

RG 93/0082

Ein Beitrag zur Kamerafokussierung bei verschiedenen Anwendungen der Bildverarbeitung

Dem Fachbereich Maschinenbau
der Universität der Bundeswehr Hamburg
zur Erlangung des akademischen Grades
eines Doktor-Ingenieurs genehmigte

DISSERTATION

von

Dipl.-Ing. Bingzi Liao
aus Wuhan (V. R. China)

DISTRIBUTION STATEMENT A

Approved for Public Release

Distribution Unlimited

Hamburg 1993

DTIC QUALITY INSPECTED A

20041112 038

RB 5733

Ein Beitrag zur Kamerafokussierung
bei verschiedenen Anwendungen
der Bildverarbeitung

Dem Fachbereich Maschinenbau
der Universität der Bundeswehr Hamburg
zur Erlangung des akademischen Grades
eines Doktor-Ingenieurs genehmigte

DISSERTATION

von
Dipl.-Ing. Bingzi Liao
aus
Wuhan (V.R. China)

Hamburg 1993

= G = 1724.

AQ F05-01-0061

Gedruckt mit Unterstützung des Deutschen Akademischen Austauschdienstes.

Tag der mündlichen Prüfung: 13. Juli 1993

Hauptreferent: Univ.-Prof. Dr.-Ing. R. Lunderstädt, UniBw H
Korreferent: Univ.-Prof. Dr.-Ing. H. Burkhardt, TU Hamburg-Harburg

Die vorliegende Arbeit entstand während meiner Tätigkeit als Stipendiat des DAAD am Institut für Automatisierungstechnik der Universität der Bundeswehr Hamburg.

Herrn Univ.-Prof. Dr.-Ing. R. Lunderstädt danke ich besonders für die Anregung und intensive Förderung der Arbeit und die Übernahme des Hauptreferates sowie für die große Hilfe während meines Aufenthaltes in der Bundesrepublik Deutschland.

Herrn Univ.-Prof. Dr.-Ing. H. Burkhardt, Technische Universität Hamburg-Harburg, danke ich für die freundliche Unterstützung und die Übernahme des Korreferates.

Ebenfalls möchte ich Frau U. Barkmann für die sorgfältige Ausführung der Schreibarbeiten sowie die geduldige Korrektur der Grammatikfehler danken.

Ungezählte Beiträge haben alle meine Kollegen des Institutes für Automatisierungstechnik beigesteuert und die Arbeit während ihrer ganzen Dauer mit Rat und Tat unterstützt. Dafür danke ich ganz herzlich.

Hamburg, im Juli 1993.

Bingzi Liao

Inhaltsverzeichnis

	<u>Seite</u>
1. Einleitung	5
1.1 Die Bedeutung der automatischen Fokussierung für die Bildverarbeitung	5
1.2 Die Geschichte der automatischen Fokussierung	7
1.3 Forschungsziele	8
2. Theoretische Grundlagen der Fokussierung	10
2.1 Das Fokussierungsprinzip des menschlichen Auges	10
2.2 Die optische Geometrie der Kameralinse	12
2.3 Die Theorie der Defokussierung	19
2.4 Die Qualitätsbewertung optischer Bilder	32
3. Automatische Fokussierungsalgorithmen	36
3.1 Anforderungen an Fokussierungsfunktionen	36
3.2 Die Varianz der Grauwerte als Fokussierungsfunktion	38
3.3 Die Entropie der Grauwerte als Fokussierungsfunktion	42
3.4 Die SMD der Grauwerte als Fokussierungsfunktion	45
3.5 Die Leistungsspektren des Bildes als Fokussierungsfunktion	47
3.6 Die Signalleistung der Grauwerte als Fokussierungsfunktion	50
3.7 "Thresholded Video Signal Pixel Count" als Fokussierungsfunktion	53
3.8 Der Laplace-Operator der Grauwerte als Fokussierungsfunktion	56
3.9 Das Differential der Varianz als Fokussierungsfunktion	57
4. Beurteilung der verschiedenen Fokussierungsfunktionen durch Experimente	60
4.1 Der Einfluß der Beleuchtung	61
4.2 Der Einfluß von Mustern des Objektes	67
4.3 Die Reaktion der Fokussierungsfunktionen bei texturfreien Bildern	69
4.4 Die Fokussierungsfunktionen bei farbigen Mustern	70
4.5 Die Fokussierungsfunktionen bei unscharfen Konturen	71
4.6 Der Einfluß durch starke Reflexion	73
4.7 Fokussierungsfunktionen bei dreidimensionalen Gebilden	75

4.8 Fokussierungsfunktionen von schwarz-weiß Mustern unterschiedlicher Grautönung	79
4.9 Zusammenfassung	81
5. Fokussierung von Bildausschnitten	87
5.1 Grundidee der Fokussierung von Bildausschnitten	87
5.2 Vorverarbeitung zur Reduzierung des Rauscheinflusses	90
5.3 Kompensierung der Zoom-Auswirkung	99
6. Fokussierung von beweglichen Objekten	103
6.1 Das Prinzip der Fokussierung von beweglichen Objekten	103
6.2 Beziehung zwischen der Objektweite und den Bildkoordinaten des Laserflecks	108
6.3 Kamerakalibrierung und Fehleranalyse	118
7. Entwurf eines Fokussierungsregelkreises	124
7.1 Max Vision AT-1 Bildverarbeitungssystem	124
7.2 Struktur von Fokussierungsregelkreisen	127
7.3 Übertragungsfunktion des Fokussierungsregelkreises	131
7.4 Reglerentwurf	136
7.5 Zeitverhalten	154
8. Zusammenfassung	162
9. Wichtige Formelzeichen	164
10. Literatur	167

1. Einleitung

1.1 Die Bedeutung der automatischen Fokussierung für die Bildverarbeitung

Die digitale Bildverarbeitung hat sich in den letzten Jahrzehnten als ein modernes und wichtiges Mittel in verschiedenen Gebieten von Wissenschaft und Technik entwickelt. Als Beispiele seien hier genannt: die Analyse und Interpretation von Luftbildern und Satellitenfotos in den Geowissenschaften, in der Meteorologie, der Astronomie und bei der Erforschung von Naturressourcen, die Analyse von optischen Röntgen-, Wärme-, radiografischen und anderen Bildern in der medizinischen Diagnostik, bei der industriellen Schadenserkenennung und in der wissenschaftlichen Forschung sowie die Analyse von Abbildungen der optischen und der akustischen Holographie.

Die Entwicklung der Hardware (Miniaturisierung der Bauteile, Erhöhung der Geschwindigkeit und der Speicherkapazität von Digitalrechnern) sowie die Reduzierung des Preises hat die Anwendungen der Bildverarbeitung in vielen Gebieten unseres Lebens verstärkt. Durch den Einsatz von modernen Bildverarbeitungssystemen wurde nicht nur eine Befreiung des Menschen von eintöniger Arbeit erreicht, sondern auch eine geringere Fehlerhäufigkeit und wesentlich höhere Verarbeitungsgeschwindigkeit erzielt.

Für die Bildverarbeitung ist es wichtig, eine Produktion von Bildern mit höchster Qualität zu erreichen. Die Bildqualität ist dabei abhängig von vielen Faktoren, z.B. der Qualität der Kamera, der Beleuchtung des Objektes und der Schärfe der Kamera. Der Fokussierungszustand der Kamera spielt die größte Rolle für die Bildqualität.

In vielen Situationen besteht der Bedarf zum Entwurf von einfachen automatischen Fokussierungsverfahren für die digitale Bildverarbeitung: dort, wo entweder das manuelle Fokussierungsverfahren nicht hinreichend empfindlich ist oder wo die automatische Fokussierung ein notwendiger Anteil des Bildverarbeitungssystems ist. Die automatische Fokussierung begünstigt maßgeblich die Forschung.

Erfahrungsgemäß wird die Fokuseinstellung mit steigenden Anforderungen an das Auflösungsvermögen immer schwieriger, d.h. die Fokuseinstellung unterliegt

starken subjektiven Einflüssen und apparativen Fehlern des Abbildungssystems, so daß sie auch bei gleichbleibendem Beobachter nur ungenau reproduzierbar ist.

Durch Anwendung der digitalen Bildverarbeitung entsteht die Möglichkeit, den Informationsgehalt eines Bildes voll zu nutzen, was bei visueller Auswertung aus physiologischen Gründen nur selten möglich ist.

Ein sehr wichtiges Problem in der allgemeinen Bildverarbeitung ist die Extraktion der Tiefeninformation aus zweidimensionalen Bildern, weil die Tiefeninformation des Objekts zur Rekonstruktion des dreidimensionalen Bildes helfen kann und schließlich zur Erkennung des Objektes und zur Interpretation der vollständigen Szene führt. Die automatischen Fokussierungsverfahren können benutzt werden, um die Tiefe des Objektes für die dreidimensionale Szenenanalyse herauszufinden.

Es gibt viele Methoden zur Ermittlung der Tiefe eines Objektes. Die traditionelle Methode ist die Anwendung der Technik von optischen oder akustischen Abstandsbestimmungsgeräten, z.B. Sonar- oder Laserabstandsbestimmungsgeräten, um den Abstand zwischen der Kamera und dem Objekt direkt zu messen. Diese Methode braucht gesteuerte Energiestrahlen (Sender) und ein Gerät zur Entdeckung der Reflexionsstrahlen (Empfänger); es ist die sogenannte direkte, aktive Methode, die keine Information von Bildern benötigt.

Die relativen und passiven Methoden zur Tiefenbestimmung basieren auf dem Bild aus einer monokularen Kamera. Die typischen Methoden sind: Texturgradient, Photometrie, Okklusionsauswirkung, Größenbeständigkeit (size constancy) und Fokussierungsverfahren.

Um die Tiefe des Objektes durch Fokussierungsverfahren zu bekommen, wird die Szene über verschiedene Fokussierungspositionen analysiert. Mit Hilfe der automatischen Fokussierung kann die beste Fokussierungsposition der Kamera für das Objekt in der Szene gefunden werden. Für eine gegebene Kameralinse wird die Tiefe des Objektes allein durch die beste Fokussierungsposition festgelegt.

1.2 Die Geschichte der automatischen Fokussierung

Die automatische Fokussierung hat man schon seit ca. 40 Jahren untersucht. Hopkins [1] hat im Jahr 1955 und Goodman [2] im Jahr 1968 die Frequenzreaktion des defokalen optischen Systems theoretisch analysiert. Hopkins hat das durch ein defokales und aberrationsfreies optische System abgebildete Bild einer Punktlichtquelle mit zwei Methoden (Wellenoptik und geometrische Optik) hergeleitet. Durch eine theoretische Analyse hat er die Toleranz für die Defekte des Fokus vom optischen System bekommen. Hopkins hat bewiesen, daß die Übertragungsfunktion eines optischen Systems durch eine Gaußverteilung annähernd beschrieben werden kann. Das defokale Bild ist gleich der Faltung des richtig fokalen Bildes mit der Punktantwort des optischen Systems. Das ist das theoretische Grundprinzip der automatischen Fokussierung.

Um die atmosphärische Verzerrung im teleskopischen Bild zu korrigieren, haben Muller und Buffington [3] im Jahr 1974 den Begriff des Schärfenindex (SI) des Bildes aufgestellt

$$SI(w) = \int_{-L}^L \int_{-R}^R \left| \frac{\partial^{n+m} f(x,y,w)}{\partial x^n \partial y^m} \right|^2 \cdot dx \cdot dy, \quad (1.1)$$

wobei f die Intensität des Punktes (x,y) in der Bildebene ist; w ist die fokale Variable des optischen Systems, die von dem Fokussierungszustand der Kamera abhängig ist; R und L sind die Dimensionen in der Bildebene; n, m sind die willkürlichen positiven Integrale. Die genannten Autoren haben den Schärfenindex $SI(w)$ im Ortsfrequenzbereich analysiert und bewiesen, daß die Funktion $SI(w)$ das Maximum in der besten fokalen Position des optischen Systems erreicht und die Veränderungsrate von $SI(w)$ in dieser Position Null ist. Im Jahr 1976 hat Erteza [4] ein automatisches Fokussierungssystem mit Hilfe des Schärfenindex realisiert. Dabei hat sich gezeigt, daß der von Muller und Buffington aufgestellte Schärfenindex für das reale Fokussierungsverfahren zu kompliziert ist.

Eine Methode zur Bestimmung des optimalen Fokus aus einer Fokuserie wurde 1971 von Boseck veröffentlicht [6]. Bei diesem Verfahren wird auf das zu prüfende Bild mit Hilfe einer lichtoptischen Diffraktionsanlage die zweidimensionale Fouriertransformation angewendet, die hierbei auf der Fraunhoferschen

Beugung beruht. Das Verfahren basiert auf theoretischen Arbeiten von Linfoot [7] und bewertet näherungsweise den Informationsgehalt des Bildes. Es ist zeitaufwendig und prinzipiell nur für fotografische Bilder anwendbar.

Ein anderes Verfahren zur Bestimmung des Optimalfokus bei einem Lichtmikroskop, das mit einer Fernsehkamera arbeitet, wurde 1976 von Jarvis [5] vorgestellt. In dieser Arbeit werden drei verschiedene Fokussierungsfunktionen verwendet und miteinander verglichen. Die von Jarvis aufgestellten Fokussierungsfunktionen sind relativ einfach und können mit Analogschaltungen realisiert werden.

Im Aufsatz "A Comparison of Different Autofocus Algorithms" haben Ligthart und Groen [8] viele andere Fokussierungsfunktionen zusammengefaßt. Grossmann [9] und Darrell [10] haben in ihren Aufsätzen über "Depth from Focus" über weitere Funktionen berichtet. Mit Hilfe von Fokussierungsverfahren haben sie die Tiefeninformation der Szene aus einem zweidimensionalen Bild erhalten.

1.3 Forschungsziele

Ziel der vorliegenden Arbeit ist es, Verfahren zu entwickeln, die geeignet sind, aus einer Bildserie des gleichen Objektes aber verschiedener Fokuspositionen der Kamera, möglichst fehlerfrei, die richtige Fokusposition der Kamera zu finden. Die Verfahren sollen den beschränkten Möglichkeiten von PC's angepaßt sein. Weiterhin sollen die Verfahren eine Echtzeit-Verarbeitung ermöglichen.

Die hier beschriebene Vorgehensweise ist in folgende Phasen geteilt:

In der ersten Phase wird die automatische Fokussierung mit acht verschiedenen Fokussierungskriterien realisiert, indem angenommen wird, daß das Objekt eine ebene Oberfläche besitzt oder die Schnittdicke des Objektes kleiner als die Schärfentiefe der Kamera ist. Das Fokussierungsprinzip aller Fokussierungskriterien wird theoretisch analysiert. Zur Beurteilung der Qualität der verschiedenen Fokussierungskriterien werden Experimente durchgeführt. Durch die gemessenen Fokussierungsfunktionen kann man die Vor- und Nachteile der einzelnen Fokussierungskriterien herausfinden. Die Qualität der Fokussierungskriterien wird nach folgenden Aspekten bewertet:

- a. Kompliziertheit der Fokussierungskriterien,
- b. Rechenzeit der Fokussierungsfunktionen,
- c. Anpassungsvermögen an Muster,
- d. Genauigkeit der Fokussierungskriterien,
- e. Zuverlässigkeit der Fokussierungskriterien.

Schließlich wird das beste Fokussierungskriterium ausgewählt.

Die zweite Phase beschäftigt sich mit der Fokussierung zu einer relativen Position des Objektes. Im Gegensatz zur ersten Phase ist die Oberfläche des Objektes jetzt nicht mehr eben. Ziel dieser Phase ist der Entwurf von Fokussierungsverfahren, mit denen die Kamera das Objekt irgendwo fokussieren kann.

Die Fokussierung von beweglichen Objekten ist die Aufgabe der dritten Phase. Wegen der Bewegung des Objektes während des Fokussierungsprozesses sind die verschiedenen bisherigen Fokussierungskriterien wirkungslos, deswegen muß ein anderes Fokussierungsverfahren gefunden werden. In der vorliegenden Arbeit wird ein zusätzliches Lasergerät benutzt. Der Abstand zwischen der Linse und dem Objekt kann mit Hilfe der Koordinaten des Laserflecks in der Bildebene direkt berechnet werden.

In der letzten Phase wird ein Fokussierungsregelkreis entworfen. Dazu wird die Struktur der Fokussierungsregelkreise zuerst analysiert und dann das mathematische Modell entwickelt. Anhand dieses mathematischen Modells wird der Regler des Systems entworfen. Schließlich wird das Zeitverhalten des geschlossenen Fokussierungskreises untersucht.

2. Theoretische Grundlagen der Fokussierung

2.1 Das Fokussierungsprinzip des menschlichen Auges

Unser subjektiver Eindruck ist so, daß wir unsere Umgebung immer als scharf ansehen, also klar fokussiert. Dieser Eindruck wird durch die praktische universale photographische Tradition, bei der das Bild immer irgendwo richtig fokal erscheint, verstärkt. Aber leider sind diese photographische Tradition und unser Eindruck der immer scharfen Fokussierungswelt nicht richtig.

Wie in Bild 2.1 gezeigt, ist das menschliche Auge nahezu kugelförmig mit einem Durchmesser von etwa 24 bis 26 mm in der knöchernen Augenhöhle freibeweglich gelagert und kann mit Hilfe dreier antagonistischer Muskelpaare bewegt werden. Unterhalb der Lederhaut liegt eine Membran (choroid) mit vielen Blutgefäßen und Pigmenten, die die Streuung des Lichtes innerhalb des Auges vermindern [11].

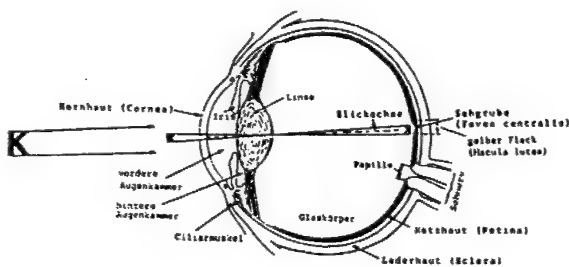


Bild 2.1: Querschnitt des menschlichen Auges.

Der vordere Teil dieser Membran wird eingeteilt in Strahlenkörper (Ciliarmuskel) und Iris, die zusammen den Blendenmechanismus bilden. Die Pupille als zentrale Öffnung der Iris hat einen variablen Durchmesser zwischen ca. 8 mm bei Dunkelheit und ca. 2 mm bei heller Lichteinwirkung. Zusätzlich zur Reaktion auf die Lichtintensität zeigt die Größe der Pupille auch verschiedene Zustände des Nervensystems an (z.B. kognitive und emotionale Belastung).

Die am weitesten nach innen liegende Membran ist die Netzhaut (Retina), auf welche bei richtiger Scharfeinstellung des Auges (Akkommodation) die Objekte der Umwelt abgebildet werden. In ihrem Zentrum liegt die Sehgrube (Fovea centralis), ein Winkelbereich von etwa $1,5^\circ$ mit der besten Auflösung.

Die abbildenden Elemente sind die Hornhaut, das Kammerwasser (zwischen Hornhaut und Linse), die Linse und der Glaskörper. Nur an der Grenzfläche zwischen Luft und Hornhaut tritt ein größerer Brechungsindexsprung von $n = 1$ zu $n = 1,38$ auf. Deshalb liegt der Hauptpunkt H (Bild 2.2) nur ca. 1,5 mm hinter dem Hornhautscheitel H_s .

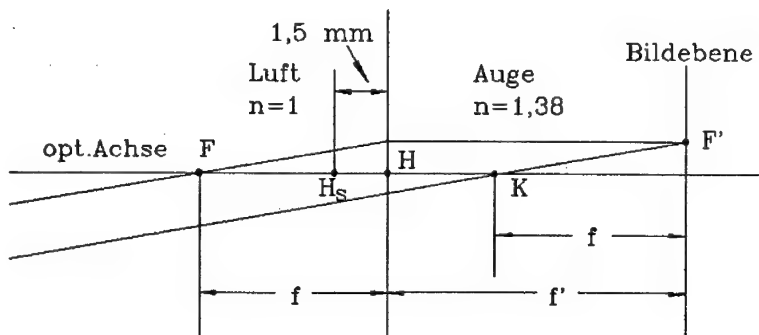


Bild 2.2: Abbildung eines Parallelstrahlbündels durch das Auge:

- F : vorderer Brennpunkt,
- F' : hinterer Brennpunkt,
- H : Hauptpunkt in der Hauptebene im Abstand $f' = 22,8$ mm von der Bildebene,
- K : Knotenpunkt im Abstand $f = 17$ mm von der Bildebene und im Abstand $5,8$ mm vom Hauptpunkt.

Die Linse ist aus konzentrischen Schichten faseriger Zellen aufgebaut und ist durch Fasern mit dem Strahlenkörper verbunden. Sie absorbiert etwa 8 % des sichtbaren Lichtes (380 - 780 nm Wellenlänge). Im Unterschied zu gewöhnlichen optischen Linsen läßt sich ihre Form ändern und zwar durch den Zug von Fasern des Ciliarmuskels.

Bei der Fokussierung auf weit entfernte Objekte (Abstand > 3 m), ergibt sich eine relativ flache Linsenform mit $f = 17$ mm. Mit kleiner werdendem Betrachtungsabstand vergrößert sich die Flächenkrümmung und f verkleinert sich auf 14 mm.

Tatsache ist, daß die auf der Netzhaut abgebildeten Bilder ganz schlecht fokal irgendwo außerhalb der Sehgrube sind. Es gibt einen Gradienten des Fokus. Der Ciliarmuskel steuert die Umdrehung des Augapfels; damit kann das Bild von dem interessanten Objekt auf die Sehgrube fallen. Das Bild außerhalb der Sehgrube ergibt sich als ein allgemeines Bild der Umgebung, ist aber unscharf. Wenn wir den Verlauf der visuellen Achse betrachten, wird deutlich, daß die Lage der Sehgrube optimal ist. Die Mitte des gerade gesehenen Bildausschnittes, d.h. was gerade fixiert wird, fällt genau auf diese Stelle. Deshalb ist die optische Achse auch der Ort, auf den sich die Aufmerksamkeit des Betrachters richtet.

Bei der Netzhaut lassen sich nun drei funktional verschiedene Schichten nachweisen, von denen wir an dieser Stelle nur den Bereich der Photorezeptoren näher betrachten wollen, Zäpfchen und Stäbchen. Diese beiden Photorezeptoren wandeln die Energie der einfallenden Lichtstrahlen durch photochemische Prozesse in bioelektrische Impulse um, die wiederum über Schaltneurone und den Sehnerv zum Gehirn gelangen und den visuellen Wahrnehmungseindruck entstehen lassen.

2.2 Die optische Geometrie der Kameralinse

Grundkenntnisse über optische Systeme (Linsen, Blenden usw.) sind unerlässlich, um die Auswirkungen der optischen Abbildung auf die notwendige und erzielbare Bildqualität des Bildverarbeitungssystems abschätzen zu können. Bild 2.3 stellt eine Kamera mit einzelner Linse dar.

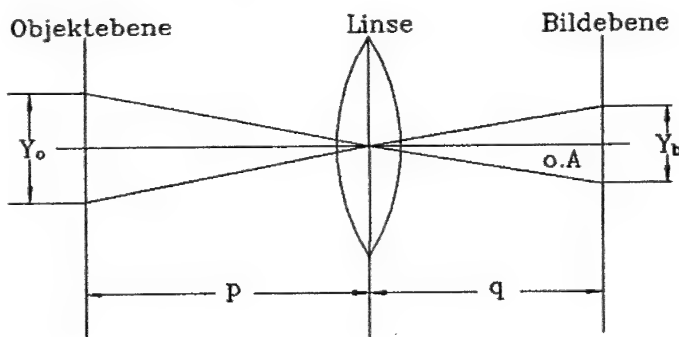


Bild 2.3: Schematische Darstellung einer Kameralinse.

Der Abstand zwischen dem Objekt und der Linse ist die Objektweite p , entsprechend ist der Abstand q zwischen der Bildebene und der Linse die Bildweite. Die Objektweite p und die Bildweite q sind mit der Brennweite f verbunden. Es gilt für die dünne Linse die bekannte Gaußsche-Linsengleichung

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f} \quad (2.1)$$

Diese Gleichung kann direkt aus der geometrischen Optik der dünnen Linse abgeleitet werden. Die Lichtstrahlen durch das Linsenzentrum beugen sich nicht, während sich die Lichtstrahlen parallel zur optischen Achse beugen, damit diese Lichtstrahlen durch den Brennpunkt gehen. Die Vergrößerung der Linse läßt sich zu

$$\beta = \frac{y_b}{y_o} \quad (2.2)$$

bestimmen. Werden die Objektweite p und die Bildweite q mit einbezogen, so ergibt sich

$$\beta = \frac{q}{p} \quad (2.3)$$

Häufig ist von Interesse zu erfahren, wie sich der Abbildungsmaßstab bei kleinen Objektverschiebungen auswirkt. Bildet man aus Gleichung (2.1)

$q = p \cdot f / (p - f)$ und differenziert nach p , so erhält man zusammen mit Gleichung (2.3) für den Tiefenmaßstab

$$\frac{dq}{dp} = -\beta^2 \quad (2.4)$$

Dieser Zusammenhang ist sehr wichtig für meßtechnische Anwendungen. Zeigt er doch, daß eine kleine Objektverschiebung in Achsenrichtung um dp zu einer um den Faktor β^2 größeren Bildverschiebung dq führt.

Theoretisch werden Linsen als lichtbrechende Medien betrachtet, ohne räumliche Begrenzungen mit einzubeziehen. Deren Berücksichtigung soll nachfolgend anhand von Blenden geschehen, die in den Strahlengang eingeführt werden. Wesentliche Vertreter hierfür sind Pupillen und Luken. Die einfachste Form einer Blende

stellt die Fassung einer Linse dar, da sie das in die Linse fallende Lichtbündel begrenzt (Bild 2.4).

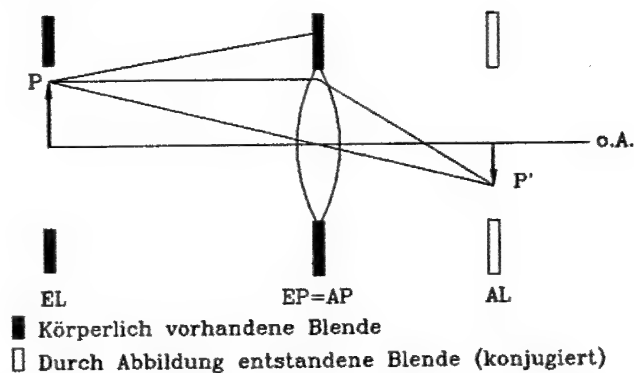


Bild 2.4: Wirkung von Blenden innerhalb eines optischen Systems.

Nicht alle vom Objektpunkt P ausgehenden Lichtstrahlen vermögen die Linse zu erreichen. Nur die Lichtstrahlen, die die Blende passieren, werden auch abgebildet. Somit wirkt diese Blende in ihren geometrischen Abmessungen als Eintrittspupille (EP). Zu ihr konjugiert, als Abbild vorliegend, ist die Austrittspupille (AP). Normalerweise sind beide gleich, d. h. $EP = AP$.

Begrenzen wir den Bereich des Objektes, von dem aus Lichtstrahlen zur Abbildung gelangen, ebenfalls durch eine Blende, in diesem Fall Eintrittsluke (EL) genannt, so wird klar, daß nur innerhalb dieser Luke liegende Objektpunkte zur Abbildung gelangen. Die Eintrittsluke führt zu einer entsprechenden Begrenzung in der Bildebene, die durch ihr Bild, die Austrittsluke (AL), gegeben ist. Eintritts- und Austrittsluke sind ebenfalls zueinander konjugiert.

Allgemein kann festgehalten werden, daß die im Strahlengang vorhandene Aperturblende als Pupille wirkt, während die Feldblende die Lukenfunktion erfüllt. Luken und Pupillen fallen niemals in einem Strahlengang zusammen. Luken haben die Funktion der Lichtfeldbegrenzung, während über die Pupillen die Helligkeit der Abbildung eingestellt werden kann.

Zur Kennzeichnung der Begrenzung von Lichtstrahlen kann die Blendenzahl BZ bzw. die numerische Apertur N.A. herangezogen werden. Aus Bild 2.5 sind diese Parameter berechenbar. Die Blendenzahl BZ ermittelt sich zu

$$BZ = \frac{f}{r} \quad , \quad (2.5)$$

wobei r der Durchmesser der Kameralinse ist. Die numerische Apertur N.A. ergibt sich allgemeiner aus dem Öffnungswinkel α zu

$$N.A. = n \cdot \sin \alpha \quad , \quad (2.6a)$$

wobei n der Brechungsindex des Mediums ist. Dies gilt sowohl für die Betrachtung auf der Objekt- als auch auf der Bildseite, so daß zwischen objekt- und bildseitiger numerischer Apertur unterschieden werden kann. Bei der Kamera ist das Medium beider Seiten Luft ($n=1$); es vereinfacht sich die Beziehung (2.6a) zu

$$N.A. = \sin \alpha \quad . \quad (2.6b)$$

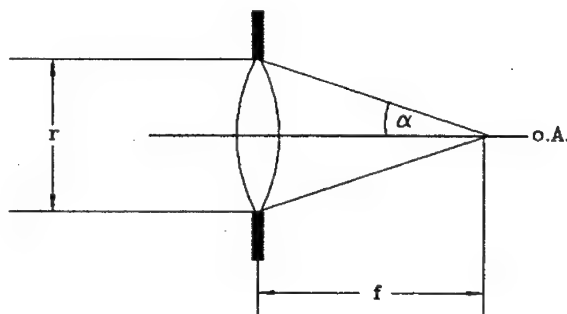


Bild 2.5: Bündelbegrenzung bei optischen Systemen.

Für die Fokussierung des Bildes spielen die Schärfentiefe S.T. und die Fokussierungstiefe F.T. eines optischen Systems eine wesentliche Rolle; S.T. beschreibt den Einfluß einer Objektverschiebung auf die Abbildungsschärfe. Die idealisierte Kamera bildet jedes Objekt "scharf" ab, unabhängig von seiner Entfernung. Bei einer realen optischen Kamera mit einem Linsensystem ist dies aus zwei Gründen nicht der Fall:

- a. Die Schärfe jeder optischen Abbildung ist aufgrund der Wellennatur des Lichtes begrenzt.

- b. Auf der Bildebene werden Objekte aus nur einer Objektebene mit bestimmter Objektweite scharf abgebildet.

Die Schärfentiefe S.T. ist der Abstand zwischen dem nächsten Punkt und dem weitesten Punkt am Objektraum entlang der optischen Achse, in der alle Objekte "zufriedenstellend" fokal sind. Eine praktische Definition des "zufriedenstellenden" Fokus kann auf dem geometrischen Modell, das in Bild 2.6 dargestellt ist, basieren.

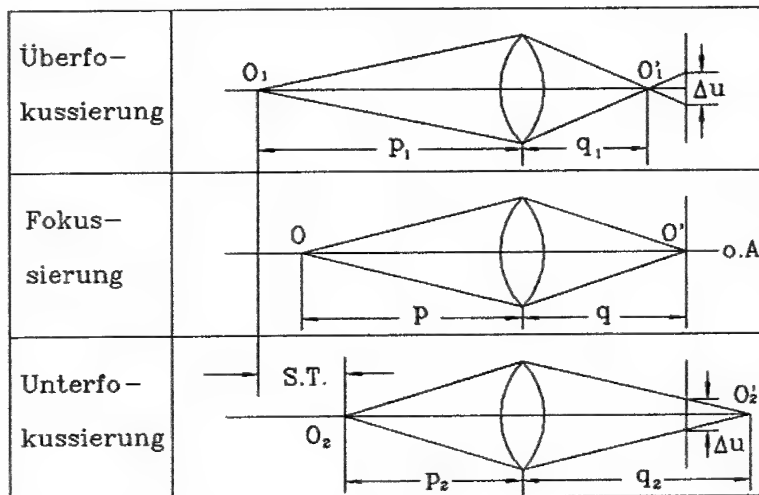


Bild 2.6: Definition der Schärfentiefe.

Eine Punktlichtquelle ist durch die Kameralinse in der Bildebene als ein Unschärfebereich abgebildet, wenn sie im defokalen Zustand ist. Der Durchmesser dieses Unschärfebereiches Δu ist ein Maß für die Defokussierung. Insbesondere kann man das Bild von der Punktlichtquelle als fokal richtig ansehen, wenn der Durchmesser Δu einen vorgegebenen Wert nicht überschreitet. Solange der Wert Δu beispielsweise kleiner als 0,1 mm auf der Bildebene bleibt, stört dies nicht, da das Auge die Punkte, die näher als 0,1 mm beieinander liegen, nicht als getrennt erkennen kann.

In Bild 2.6 wird eine Punktlichtquelle, die zwischen p_1 und p_2 an der optischen Achse liegt, zwischen q_1 und q_2 im Bildraum fokal abgebildet. In dieser Tiefe ist der Durchmesser des Unschärfebereiches kleiner als 0,1 mm; p_1 und p_2 definieren die weiteste und nächste Objektebene.

p_1 und p_2 können als Funktionen von p , f , Δu und dem Durchmesser der Linse r ausgedrückt werden. Die Herleitung basiert auf einfachen geometrischen Überlegungen. Aus der Ähnlichkeit des Dreiecks im Bildraum ergibt sich

$$\frac{\Delta u}{r} = \frac{q - q_1}{q_1} = \frac{q_2 - q}{q_2} \quad (2.7)$$

Setzt man (2.1) in (2.7) ein, so folgt

$$\frac{\Delta u}{r} = \frac{p_1 \cdot f - p \cdot f}{p_1 \cdot p - p_1 \cdot f} = \frac{p \cdot f - p_2 \cdot f}{p_2 \cdot p - p_2 \cdot f} \quad (2.8)$$

Aus Gleichung (2.8) kann man dann p_1 und p_2 berechnen. Man erhält

$$p_1 = \frac{p \cdot r \cdot f}{r \cdot f - \Delta u \cdot (p - f)} \quad (2.9a)$$

$$p_2 = \frac{p \cdot r \cdot f}{r \cdot f + \Delta u \cdot (p - f)} \quad (2.9b)$$

Die gewünschten Ergebnisse folgen somit unmittelbar:

$$\text{weiteste Tiefe } D1 = p_1 - p = \frac{\Delta u \cdot p \cdot (p - f)}{r \cdot f - \Delta u \cdot (p - f)} \quad (2.10a)$$

$$\text{nächste Tiefe } D2 = p - p_2 = \frac{\Delta u \cdot p \cdot (p - f)}{r \cdot f + \Delta u \cdot (p - f)} \quad (2.10b)$$

$$\text{S.T.} = D1 + D2 = \frac{2 \cdot p \cdot r \cdot f \cdot \Delta u \cdot (p - f)}{r^2 \cdot f^2 - \Delta u^2 \cdot (p - f)^2} \quad (2.10c)$$

Die Bedeutung dieser Gleichungen für unsere Zwecke ist:

- a) Die Schärfentiefe S.T. wird in dem Maße abnehmen wie der Durchmesser der Linse zunimmt. Es kommt darin der anschauliche Sachverhalt zum Ausdruck, daß die Schärfentiefe mit kleiner werdendem Durchmesser der Linsenapertur anwächst. Um die Schärfentiefe des optischen Systems zu erhöhen, kann man den Durchmesser der Linse verkleinern, aber eine Kamera mit zu kleiner Linsenapertur absorbiert nicht genug Strahlen. Deshalb ist die Auswahl des Durchmessers der Linse ein Kompromiß zwischen der Schärfentiefe und der Fähigkeit zur Absorbierung der Energie. Die Apertur des menschlichen

Auges ist zehnmal kleiner als die einer guten Kamera, entsprechend ist die Schärfentiefe zehnmal größer.

- b) Die Schärfentiefe S.T. nimmt ab, wenn die Brennweite der Linse zunimmt. Diese Tatsache kann man durch die Umschreibung der Linsengleichung einfach verstehen

$$(p-f) \cdot (q-f) = f^2 \quad (2.11)$$

Diese Gleichung ist in Bild 2.7 gezeichnet.

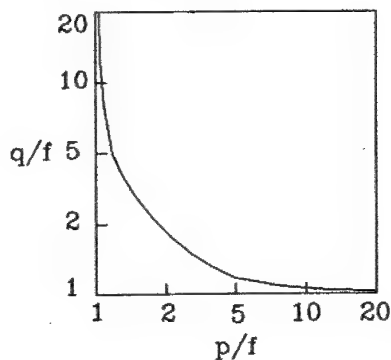


Bild 2.7: Schematische Darstellung der Linsengleichung.

In diesem Bild sind beide Achsen auf f normiert. Die Bildweiten q_1 und q_2 entsprechen einem willkürlichen Intervall der Objektweite in der Objektebene, z.B. 10 mm und 20 mm. Mit einer 1 mm-Linse ($f=1$) liegen diese Tiefen an der horizontalen Asymptote des Bildes. So besteht ein unbedeutender Unterschied zwischen der q -Koordinate und ihrer jeweilig besten Bildebene. Aus dieser Situation ergibt sich eine breite Schärfentiefe. Nun wird die Auswirkung einer 4 mm-Linse betrachtet. Die zu fokussierenden Punkte liegen ebenfalls zwischen 10 mm und 20 mm. Diese beiden Tiefen befinden sich an der normierten Achse bei 2,5 bzw. 5 und produzieren einen sehr großen Unterschied an q -Koordinaten (weil die q -Achse auch durch f normiert ist). Daraus resultiert, daß eine gleichzeitige Fokussierung dieser beiden Punkte sehr schwer ist. Die Fokussierungstiefe F.T. führt die gestattete Toleranz in der Bildebene, bei der das Bild eines Objektes mit bestimmter Objektweite in einem akzeptablen Fokus bleibt, zurück. In Bild 2.8 ist der Abstand F.T. der Fokussierungstiefe für einen gestatteten Durchmesser des Unschärfebereichs Δu gezeichnet.

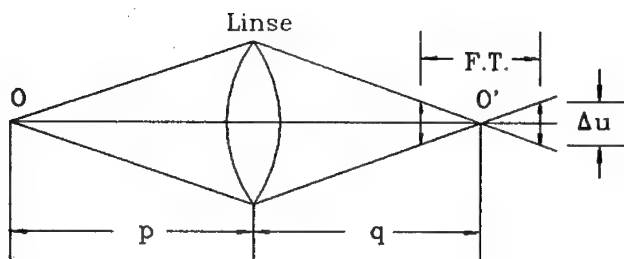


Bild 2.8: Definition der Fokussierungstiefe.

Aus dem Bild kann man ersehen, daß dieses Intervall symmetrisch zur besten Fokussierungsebene liegt. Es ist

$$F.T. = \frac{2 \cdot q \cdot \Delta u}{r} \quad (2.12a)$$

Für $p \gg f$, also $q \approx f$, ergibt sich

$$F.T. \approx \frac{2 \cdot f \cdot \Delta u}{r} \quad (2.12b)$$

Die Fokussierungstiefe mißt, wie scharf die beste Bildebene im Bildraum definiert ist. Sie ist ein brauchbares Kriterium zur Festlegung der Genauigkeit von automatischen Fokussierungskriterien. Durch die o.g. Analysen kann man sehen, daß die beste Fokussierung durch die Veränderung der Objektweite oder der Bildweite erreicht wird.

2.3 Die Theorie der Defokussierung

Der Begriff Bild bezieht sich auf eine zweidimensionale Funktion der Lichtintensität und ist mit $f(x,y)$ oder $g(x,y)$ bezeichnet. Der Wert $f(x,y)$ für die Ortskoordinaten (x,y) bezeichnet die Intensität (Helligkeit) des Bildes an diesem Punkt. Weil das Licht eine Form von Energie ist, muß $f(x,y)$ zwischen Null und Unendlich liegen, also

$$0 < f(x,y) < \infty \quad (2.13)$$

Das Bild, das wir jeden Tag visuell wahrnehmen, besteht aus dem Reflexionslicht des Objektes. Der Wertebereich der Bildfunktion $f(x,y)$ ergibt sich aus dem Produkt der Beleuchtungsstärke $i(x,y)$ und des Reflexionsgrades der aufgenommenen Objekte $l(x,y)$

$$f(x,y) = i(x,y) \cdot l(x,y) \quad , \quad (2.14)$$

wobei

$$0 < i(x,y) < \infty \quad (2.15a)$$

und

$$0 < l(x,y) < 1 \quad (2.15b)$$

sind.

Die Gleichung (2.15b) bedeutet, daß der Reflexionsgrad $l(x,y)$ durch 0 (totale Absorption) und 1 (totale Reflexion) begrenzt ist. Die Funktion $i(x,y)$ ist durch die Beleuchtungsquelle bestimmt, und $l(x,y)$ hängt von den Eigenschaften des Objektes in der Szene ab. Die Intensität eines monochromen Bildes an den Koordinaten (x,y) wird als Grauwert dieses Bildes an diesem Punkt bezeichnet.

Für eine bestimmte Szene sind die maximalen und minimalen Objektleuchtedichten L_{\max} und L_{\min} proportional zur oberen und unteren Grenze der Grauwerte. Mit a als Proportionalitätsfaktor ist

$$0 \leq a \cdot L_{\min} \leq f(x,y) \leq a \cdot L_{\max} \quad . \quad (2.16)$$

Diese Grenzen bestimmen den Grauwertbereich. Üblicherweise transformiert man diesen Bereich in einen Bereich $[0,A]$, wobei schwarz gleich Null und weiß gleich A angenommen wird. Für die Verarbeitung im Digitalrechner müssen die unabhängigen Variablen (x,y) und die Grauwerte $f(x,y)$ digitalisiert werden. Nach einer äquidistanten Digitalisierung in L Zeilen und R Spalten erhält man eine Matrix

$$E = \begin{bmatrix} f(x,y) \end{bmatrix} \quad \text{mit} \quad \begin{array}{l} x = 0, 1, 2, \dots, L-1, \\ \text{und} \quad y = 0, 1, 2, \dots, R-1, \end{array}$$

deren einzelne Elemente als Bildpunkte bezeichnet werden. Für die entsprechenden quantisierten Grauwerte gilt

$$0 < f(x,y) < G-1 \quad .$$

Es ist günstig, die Zahl der Quantisierungsstufen als ganzzahlige Potenzen der Zahl zwei zu wählen

$$L = R = 2^n \quad \text{und} \quad G = 2^m$$

Die Qualität eines Schwarzweißbildes wird in guter Näherung durch ein entsprechendes 512 x 512 Digitalbild mit 256 Graustufen erreicht, d.h. $n = 9$ und $m = 8$.

Die optische Abbildung wird als Prozeß einer Signalübertragung aufgefaßt, bei der das optische (abbildende) System die Rolle eines Frequenzfilters spielt. Diese Beschreibungsweise ist eine vollständige Analogie zur Übertragung elektrischer Nachrichten durch Frequenzfilter, wo die Nachricht eine Funktion nur der Zeit und nicht des Ortes ist.

Die optische Abbildung als ein allgemeines zweidimensionales System H (Bild 2.9) bildet ein Eingangsbild $f(x,y)$ eindeutig auf ein Ausgangsbild ab

$$g(x,y) = H\{f(x,y)\} \quad (2.17)$$

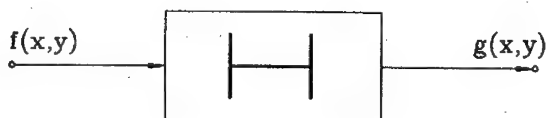


Bild 2.9: Allgemeines 2-D System.

Besondere Bedeutung haben lineare Systeme. Das System H ist genau dann linear, wenn bei jeder Linearkombination von zwei Eingangsbildern $f_1(x,y)$ und $f_2(x,y)$ die Systemreaktion des Summenbildes darstellbar ist als dieselbe Linearkombination der Systemreaktion auf die Einzelbilder, d.h., wenn

$$H(a_1 \cdot f_1(x,y) + a_2 \cdot f_2(x,y)) = a_1 \cdot H\{f_1(x,y)\} + a_2 \cdot H\{f_2(x,y)\} \quad (2.18)$$

gilt, wobei a_1 und a_2 Konstanten sind.

Im weiteren sei die δ -Funktion $\delta(x-x', y-y')$ eine Punktlichtquelle in der Objektebene am Ort $(x, y) = (x', y')$. Aus dieser δ -Funktion, die jeweils linear unabhängige Punktlichtquellen darstellen kann, lassen sich durch lineare Superposition beliebig geformte Eingangsbilder aufbauen, also

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x', y') \cdot \delta(x-x', y-y') \cdot dx' \cdot dy' \quad , \quad (2.19)$$

wobei $f(x', y') \cdot dx' \cdot dy'$ als Gewichtsfunktion für die δ -Lichtquellen aufgefaßt werden kann. Das der Funktion $f(x, y)$ entsprechende Eingangsbild ergibt sich durch lineare Superposition der Bilder der einzelnen Punktlichtquellen im Objektraum.

Die Linearitätsbedingung (2.18) bedeutet im Falle der Abbildung einer Punktlichtquelle in ein Punktbild (Punktantwort), daß eine Veränderung des Eingangsbildes um den Faktor a eine Veränderung des Ausgangsbildes um den gleichen Faktor hervorruft. Darüber hinaus wird die Übertragung dieses Bildes nicht gestört durch die gleichzeitige Abbildung anderer Objektpunkte, d.h. es gibt keine Interaktion der verschiedenen parallelen Übertragungswege. Bei rein optischen Abbildungssystemen sind diese Bedingungen für vollständige Inkohärenz oder Kohärenz der Wellen immer erfüllt. Es ergibt sich das Ausgangsbild des optischen Systems gemäß Bild 2.9 zu

$$g(x, y) = H \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') \cdot \delta(x-x', y-y') \cdot dx' \cdot dy' \right\} \quad , \quad (2.20a)$$

oder

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') \cdot H\{\delta(x-x', y-y')\} \cdot dx' \cdot dy' \quad . \quad (2.20b)$$

Der zweite Faktor von (2.20b), der redifiniert ist als

$$h(x, y, x', y') = H\{\delta(x-x', y-y')\} \quad , \quad (2.21)$$

ist die Punktantwort des optischen Systems. Setzt man die Punktantwort in (2.20b) ein, so ergibt sich

$$g(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x',y') \cdot h(x,y,x',y') \cdot dx' \cdot dy' \quad . \quad (2.22)$$

Neben der Linearität wurde noch eine weitere Eigenschaft an das optische System gestellt, nämlich die Verschiebungsinvarianz der Punktantwort. Um diesen Begriff zu erklären, wird eine Punktlichtquelle im Ursprung betrachtet, also $f(x,y) = \delta(0,0)$. Aus der angenommenen Rotationssymmetrie des optischen Abbildungssystems folgt, daß auch die Punktantwort $h(x',y')$ Rotationssymmetrie besitzen muß. Verschiebt man die Punktlichtquelle an eine Stelle (x',y') unterschiedlich von $(0,0)$, so verschiebt sich die Punktantwort ebenfalls in der Bildebene. Durch die Bedingung, daß sich die Gestalt der Punktantwort durch Verschiebung nicht ändert, ist die Punktantwort nur von der Differenz der Koordinaten $x - x'$ und $y - y'$ abhängig, also

$$h(x,y,x',y') = h(x-x',y-y') \quad . \quad (2.23)$$

In der Praxis sind die Linearität und die Verschiebungsinvarianz des optischen Systems in globaler Form nicht erfüllbar, z.B. ist die Punktantwort nie über der ganzen Bildebene stationär. Da sie sich aber als Funktion des Ortes nur langsam (stetig) ändert, sind sogenannte Isoplanasiebezirke angebbbar, in denen die Verschiebungsinvarianz näherungsweise Geltung hat. Die Eigenschaften der optischen Übertragung werden allerdings damit zu einer Funktion des untersuchten Isoplanasiebezirkes.

Gegenüber der Linearität sind ebenfalls ernste Vorbehalte anzubringen. Die lineare Überlagerung von Intensitäten setzt nämlich die völlige Inkohärenz der beteiligten Lichtwellen voraus.

Wenn das optische System die Bedingungen (2.18) und (2.23) erfüllt, dann folgt die Beziehung

$$\begin{aligned} g(x,y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x',y') \cdot h(x-x',y-y') \cdot dx' \cdot dy' \\ &= f(x,y) * h(x,y) \quad . \end{aligned} \quad (2.24)$$

Das Superpositionsintegral (2.22) verwandelt sich zu der Faltungsoperation (2.24).

Ganz analog zur Fouriertransformation einer nicht periodischen eindimensionalen Zeitfunktion wird die Fouriertransformation eines Bildes $f(x,y)$ mit den beiden unabhängigen Ortskoordinaten x und y definiert durch

$$F(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \cdot \exp(-2\pi \cdot j \cdot (u \cdot x + v \cdot y)) \cdot dx \cdot dy \quad (2.25)$$

Im allgemeinen ist $F(u,v)$ eine komplexe Funktion von Ortsfrequenzen u und v mit einem Realteil $F_r(u,v)$ und einem Imaginärteil $F_i(u,v)$:

$$F(u,v) = F_r(u,v) + j \cdot F_i(u,v) = |F(u,v)| \cdot \exp(j \cdot \Phi(u,v)) \quad ,$$

mit

$$|F(u,v)| = \sqrt{[F_r^2(u,v) + F_i^2(u,v)]}$$

und

$$\Phi(u,v) = \tan^{-1} \left[\frac{F_i(u,v)}{F_r(u,v)} \right] \quad .$$

$F(u,v)$ heißt Frequenzspektrum von $f(x,y)$ und $|F(u,v)|$ das Fourierspektrum von $f(x,y)$; $\Phi(u,v)$ ist der Phasenwinkel von $f(x,y)$, $|F(u,v)|^2 = F_r^2(u,v) + F_i^2(u,v)$ ist das Leistungsspektrum des Bildes.

Wegen der Eigenschaften der Fouriertransformation ergibt sich die Fouriertransformation des Faltungsintegrals (2.24) zu

$$G(u,v) = F(u,v) \cdot H(u,v) \quad , \quad (2.26)$$

mit

$$G(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \cdot \exp(-2\pi \cdot j \cdot (u \cdot x + v \cdot y)) \cdot dx \cdot dy$$

und

$$H(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) \cdot \exp(-2\pi \cdot j \cdot (u \cdot x + v \cdot y)) \cdot dx \cdot dy \quad .$$

Die Fouriertransformation der Punktantwort $H(u,v)$ wird als die optische Übertragungsfunktion (ÖÜF) bezeichnet. Der Betrag $|H(u,v)|$ ist als Modulationsübertragungsfunktion (MÜF) (englisch: Modulation Transfer Function) bekannt.

Um einen Zusammenhang zwischen den Eigenschaften eines optischen Systems und seiner ÖÜF zu finden, wird vorerst die Punktantwort bestimmt. Die Abbildung des optischen Systems wird dazu zunächst vom Standpunkt der Wellenoptik aus behandelt. Einfach gesagt, die Defokussierung eines inkohärenten optischen Systems verursacht die Unschärfe des Bildes. Es ist möglich, daß die Defokussierung als eine Wellenaberration behandelt wird. Diese Aberration wird durch die optische Übertragungsfunktion ÖÜF dargestellt. Für die Herleitung werden reduzierte Koordinaten gemäß Bild 2.10 benutzt.

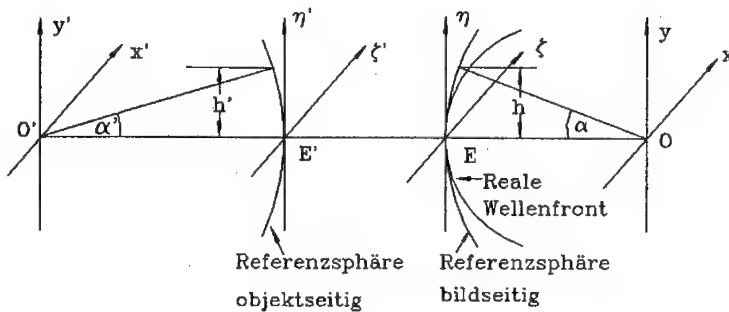


Bild 2.10: Schematische Darstellung des optischen Systems,
 O', O : Objekt- und Bildebene,
 E', E : Eintritts- und Austrittspupille.

Die kleinen Buchstaben sind reduzierte Koordinaten, die zu den entsprechenden, durch Großbuchstaben gekennzeichneten, geometrischen Koordinaten gehören. Für Bild- und Objektebene definiert man

$$\begin{aligned} x &= \frac{n \cdot \sin \alpha}{\lambda} \cdot X, & x' &= \frac{n' \cdot \sin \alpha'}{\lambda} \cdot X', \\ y &= \frac{n \cdot \sin \alpha}{\lambda} \cdot Y, & y' &= \frac{n' \cdot \sin \alpha'}{\lambda} \cdot Y', \end{aligned} \quad (2.27)$$

mit den numerischen Aperturen $n \cdot \sin \alpha$ bild- und $n' \cdot \sin \alpha'$ objektseitig. Mit den Pupillenhöhen h und h' definiert man weiter die reduzierten Pupillenkoordinaten

$$\zeta = \frac{Z}{h}, \quad \eta = \frac{H}{h}, \quad \zeta' = \frac{Z'}{h'}, \quad \eta' = \frac{H'}{h'} \quad (2.28)$$

Es folgt $\zeta = \zeta'$ und $\eta = \eta'$ bei Vernachlässigung der Verzeichnung.

In Eintritts- und Austrittspupille werden, wie in Bild 2.10 gezeigt, Referenzsphären mit den Zentren O' und O gelegt. Die Wellenfront (= Fläche konstanter Phase) einer von O' ausgehenden Lichtwelle hat im Falle eines idealen Systems im Bildraum wieder die Form einer Kugelwelle, die in den Punkt O einläuft. Bei Systemen mit Aberration treten Abweichungen $W(\zeta, \eta)$ von der Kugelform auf, entsprechend bei den auf die Austrittsreferenzsphäre bezogenen Phasen $k \cdot W(\zeta, \eta)$. Mit den Amplituden der Lichterregung $a(\zeta, \eta)$ kann man die der Referenzsphäre zugeordnete Pupillenfunktion

$$b(\zeta, \eta) = a(\zeta, \eta) \cdot \exp(j \cdot k \cdot W(\zeta, \eta)) \quad (2.29)$$

definieren. Sie hat die Pupillenfläche $\zeta^2 + \eta^2 \leq 1$ oder $Z^2 + H^2 \leq h^2$ als Träger. Für die praktische Anwendung in der OÜF-Berechnung betrachtet man den Lichtfluß durch die Austrittspupille als homogen und setzt

$$a(\zeta, \eta) = \begin{cases} 1 & \text{für } \zeta^2 + \eta^2 \leq 1 \\ 0 & \text{sonst} \end{cases} \quad (2.30)$$

Die Kenntnis der Pupillenfunktion $b(\zeta, \eta)$ genügt, um die Lichtamplitudenverteilung $h_1(x, y)$ in der Bildebene zu ermitteln (Bild 2.11).

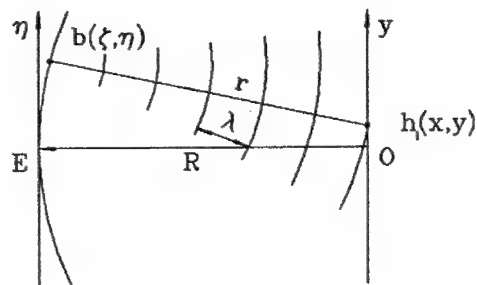


Bild 2.11: Pupillenfunktion.

Jeder Punkt auf der Referenzsphäre kann als Ausgangspunkt einer Kugelwelle

$$\frac{1}{r} \cdot \exp(-j \cdot k \cdot r)$$

betrachtet werden. Man erhält das Beugungsintegral

$$h_1(x, y) = c \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\xi, \eta) \cdot \frac{\exp(-j \cdot k \cdot r(\xi, \eta, x, y))}{r(\xi, \eta, x, y)} \cdot d\xi \cdot d\eta \quad , \quad (2.31)$$

wobei c eine Konstante ist. Für das kleine Gebiet in der Gegend der optischen Achse, in dem $h_1(x, y)$ wesentlich von 0 verschieden ist, kann der Nenner als konstant angesehen und zur Konstanten hinzugeschlagen werden. Der Exponent ist entwickelbar und man erhält

$$r \approx R - \frac{2 \cdot \pi}{k} \cdot (\xi \cdot x + \eta \cdot y) \quad (2.32)$$

in reduzierten Koordinaten mit $k = 2 \cdot \pi \cdot n / \lambda$ (Wellenzahl im Bildraum). Setzt man (2.32) in (2.31) ein, erhält man das bekannte Fraunhofer'sche Beugungsintegral

$$h_1(x, y) = c_1 \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\xi, \eta) \cdot \exp(+j \cdot 2\pi \cdot (\xi \cdot x + \eta \cdot y)) \cdot d\xi \cdot d\eta \quad , \quad (2.33)$$

wobei $c_1 = c \cdot \exp(-j \cdot R)$ ist, und weiter die Intensität $h(x, y)$ der Punktantwort

$$h(x, y) = h_1(x, y) \cdot h_1^*(x, y) \quad . \quad (2.34)$$

Die berechnete Amplitudenverteilung (2.33) kann nach der Pupillenfunktion aufgelöst werden und man erhält

$$b(\xi, \eta) = c_2 \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(x, y) \cdot \exp(-j \cdot 2\pi \cdot (x \cdot \xi + y \cdot \eta)) \cdot dx \cdot dy \quad , \quad (2.35)$$

wobei c_2 eine komplexe Konstante ist.

Die Pupillenfunktion ist also proportional zur Fouriertransformation der Amplitudenverteilung $h_1(x, y)$ in der Punktantwort. Ferner ist wegen (2.34) die

OÜF proportional zur Fouriertransformation der Intensitätsverteilung $h(x,y)$.
Mit dem Faltungstheorem wird

$$H(u,v) = c \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\zeta, \eta) \cdot b^*(\zeta - u, \eta - v) \cdot d\zeta \cdot d\eta \quad , \quad (2.36)$$

wobei u und v wieder geeignete Ortsfrequenzen sind.

Wegen der endlichen Träger von $b(\zeta, \eta)$ ist eine Verschiebung des Koordinatennetzes zulässig. Man erhält so die symmetrische Form der obigen Gleichung

$$H(u,v) = c \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\zeta + \frac{u}{2}, \eta + \frac{v}{2}) \cdot b^*(\zeta - \frac{u}{2}, \eta - \frac{v}{2}) \cdot d\zeta \cdot d\eta \quad . \quad (2.37)$$

Die Gleichung (2.37) ist das sogenannte Duffieux-Integral; sie gibt den Zusammenhang zwischen der Pupillenfunktion und der OÜF an. Die OÜF ist die Autokorrelation der Pupillenfunktion. Die Integration in (2.37) ist nur scheinbar über einem unendlichen Gebiet durchzuführen, praktisch erstreckt sie sich nur über den Durchschnitt S der um $\pm (u/2, v/2)$ gescherten Träger A der Pupillenfunktion (Bild 2.12).

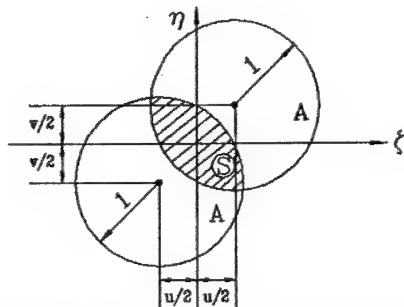


Bild 2.12: Region der Integration für die Ortsfrequenzen (u,v) .

Für ein aberrationsfreies optisches System in der Anwesenheit einer Defokussierung hat die Pupillenfunktion die Form

$$b(\zeta, \eta) = \begin{cases} \exp(j \cdot k' \cdot w_{20} \cdot (\zeta^2 + \eta^2)) & \text{für } \zeta^2 + \eta^2 \leq 1 \\ 0 & \text{sonst} \end{cases} \quad . \quad (2.38)$$

k' ist eine Konstante, w_{20} ist eine Funktion von α und Δq .

Wie Bild 2.13 zeigt, ist die Funktion w_{20} die optische Weglänge des Abschnitts zwischen der realen Wellenfront und der Referenzsphäre. Das Zentrum der realen Wellenfront ist die fokale Bildebene O_0 . Das Zentrum der Referenzsphäre ist die reale Bildebene O . w_{20} stellt den Fokussierungsdefekt des optischen Systems dar. Wenn das optische System richtig fokussiert, überlappen sich die Referenzsphäre und die Wellenfront. Wenn zwischen der realen Bildebene und der fokalen Bildebene eine Verschiebung Δq auftritt, ist der Fokussierungsdefekt w_{20} durch folgende Gleichung gegeben

$$w_{20} = \frac{1}{2} \cdot n \cdot \sin^2 \alpha \cdot \Delta q \quad (2.39)$$

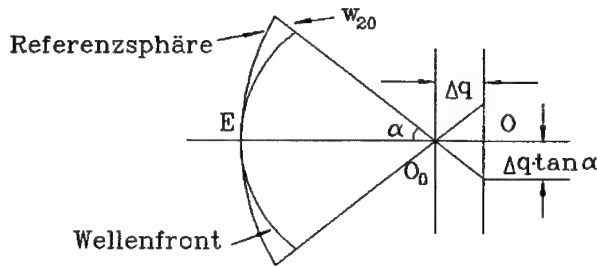


Bild 2.13: Messung des Fokussierungsdefektes.

Wenn man die Pupillenfunktion (2.38) in die Gleichung (2.36) einsetzt, ist die Frequenzreaktion direkt durch den Grad der Defokussierung gegeben. Die OÜF des optischen Systems kann als eine Potenzreihe in Bessel-Funktionen dargestellt werden. Die durch die Wellenoptik erhaltene OÜF ist sehr kompliziert. In der Praxis wird die OÜF aus der geometrischen Optik heraus abgeleitet.

Gemäß der geometrischen Optik ist die Punktantwort innerhalb des Unschärfebereichs annähernd konstant [12]

$$h(x,y) = \begin{cases} \frac{4}{\pi \cdot \Delta u^2} & \text{für } x^2 + y^2 \leq \frac{\Delta u^2}{4} \\ 0 & \text{sonst} \end{cases} \quad (2.40a)$$

Wegen der Diffraktion, der Aberration der Linse und anderen Auswirkungen wird die Punktantwort besser als eine zweidimensionale Gauß-Verteilung dargestellt

$$h(x,y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (2.40b)$$

Im mathematischen Sinn ist σ die Streuung der Gauß-Verteilung. Hier ist σ aber der Ausdehnungsparameter des optischen Systems, nämlich

$$\sigma = k \cdot \Delta u \quad (2.41a)$$

k ist eine positive Konstante, die von der Eigenschaft des optischen Systems abhängig ist. Durch eine einfache Rechnung kann man Δu bestimmen. Man erhält

$$\Delta u = r \cdot q \cdot \left(\frac{1}{f} - \frac{1}{p} - \frac{1}{q}\right) \quad (2.41b)$$

Von (2.40b) kann die OÜF des optischen Systems hergeleitet werden. Es folgt

$$H(u,v) = \exp\left(-\frac{1}{2} \cdot (u^2+v^2) \cdot \sigma^2\right) \quad (2.42)$$

Die OÜF $H(u,v)$ kann als ein Tiefpaßfilter behandelt werden. In Polarkoordinaten ρ, ψ ergibt sich aus (2.42)

$$H(\rho, \psi) = \exp\left(-\frac{1}{2} \cdot \rho^2 \cdot \sigma^2\right) \quad (2.43)$$

mit

$$\begin{aligned} \rho &= \sqrt{u^2 + v^2} \quad , \\ \psi &= \tan^{-1} \frac{u}{v} \quad . \end{aligned}$$

ρ ist die Ortsfrequenz in Polarkoordinaten. Der Ausdehnungsparameter σ ist proportional zum Grad der Defokussierung. Bei guter Fokussierung besitzt das optische System ein kleines σ , im Gegensatz dazu hat ein reales optisches System ein großes σ .

In Bild 2.14 kann man sehen, daß, wenn $\sigma = 0$ ist, die OÜF wie ein Allpaßfilter arbeitet. Das aufgenommene Bild ist ein perfekt fokales Bild. Wenn σ zunimmt, wird die OÜF im hohen Frequenzbereich schnell abnehmen. Das Bild ist durch die Defokussierung unscharf geworden.

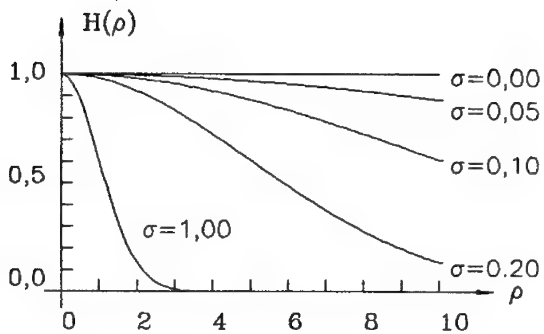


Bild 2.14: Schematische Darstellung der OÜF.

Für ein lineares verschiebungsinvariantes System ist das Ausgangsbild $g(x,y)$ durch die Faltung des Eingangsbildes mit der Punktantwort gegeben. Das Eingangsbild kann man als ein richtig fokales Bild behandeln und das Ausgangsbild als ein defokales Bild.

Bild 2.15 zeigt die Auswirkung der Defokussierung am Profil einer idealen Kante. Der Übergang der defokalen Kante ist Δu .

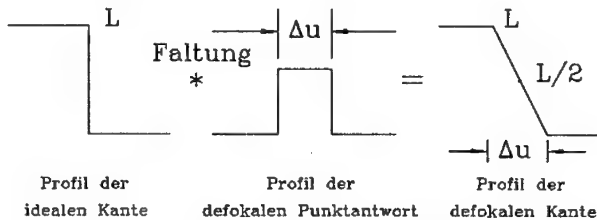


Bild 2.15: Auswirkung der Defokussierung.

Durch die obige Analyse wird deutlich, daß der Unterschied zwischen dem fokalen und defokalen Bild im hohen Frequenzbereich liegt. Die Leistungsspektren im hohen Frequenzbereich sind deshalb ein Maß für die Defokussierung.

2.4 Die Qualitätsbewertung optischer Bilder

Die Produktion von Bildern mit höchster Qualität ist die natürliche Anforderung, die ein Fotograf stellt. Die Bildqualität ist aber eine subjektive Schätzung. Es ist unmöglich, daß die Bildqualität als eine einfache Funktion der physikalischen Variablen betrachtet wird. Weil die Struktur des photographischen Bildes durch verschiedene physikalische Variablen gekennzeichnet ist, ist die Vorhersage der Bildqualität ohne Einschränkungen unmöglich. L.M. Bibermann [12] hat das Problem gut formuliert.

"In general the concept of image quality does not emerge until we have defined some tasks for the system to perform; such tasks do not necessarily have to be performed by the components. Definition of the imaging task is a vital part of image evaluation. The more limited and precisely defined we can make the task, the more accurately can the performance of the system be evaluated; the converse is also true."

Für den Zweck der automatischen Fokussierung wird angenommen, daß der einzige Faktor zur Beeinflussung der Bildqualität die Schärfe des optischen Systems ist.

U. Pfenninger [13] hat festgestellt, daß die Beurteilungswerte der Bildqualität und der Schärfe des Bildes eine Linearkombination bilden

$$\psi_q = a_0 + a_s \cdot \psi_s \quad , \quad (2.44)$$

wobei ψ_q der Skalierungswert für die Bildqualität ist und ψ_s der Skalierungswert für die Schärfe des Bildes, a_0 und a_s sind zwei Konstanten. Das bedeutet, daß die Bildqualität linear abhängig von der Schärfe des Bildes ist. Normalerweise ist die Schärf beurteilung eine Funktion der OÜF, und zwar

$$\psi_s = K_s \cdot \frac{\log H(\rho)}{(-c \cdot \rho)} + B_s \quad . \quad (2.45)$$

K_s , B_s und c sind Konstanten. Die Gleichung (2.45) gilt für die Bedingung, daß die OÜF eine Symmetrie zum Ursprung der Ortsebene besitzt. In diesem Fall genügt zur Beschreibung die Angabe der ebenfalls rotationssymmetrischen OÜF mit einer einzigen Ortsfrequenzkoordinate bzw. die Beschreibung in Polarkoordinaten.

Viele Kriterien der Bildqualität basieren auf der Übertragungsfunktion des optischen Systems und des menschlichen Auges, weil die Qualität des vom optischen System aufgenommenen Bildes durch das menschliche Auge beurteilt wird.

a. SMT (System Modulation Transfer Acutance)-Schärfe

Die älteste Messung der Schärfe durch die OÜF wurde von E.M. Grane [8] schon im Jahr 1964 vorgeschlagen. Die Messung der Bildqualität basiert auf der Hypothese, daß alle Punktantworten der Gaußverteilung genügen. Die OÜF ist in diesem Fall auch die Gaußverteilung. Die Fläche unter der Funktion $H(\rho)$ ist proportional zu $1/\sigma^2$. Zusammen mit dem Ausdehnungsparameter des visuellen Systems σ_R kann man eine Messung der Bildqualität bekommen. Die sogenannte SMT-Schärfe ist dann definiert als

$$\text{SMT-Schärfe} = K + k \cdot \log\left(\frac{\sigma_R}{\sigma}\right) \quad , \quad (2.46)$$

wobei K, k Konstanten sind.

b. CMT (Cascaded Modulation Transfer Function)-Schärfe

Die CMT-Schärfe von R.G. Gendron [15] ist in Bild 2.16 gezeigt.

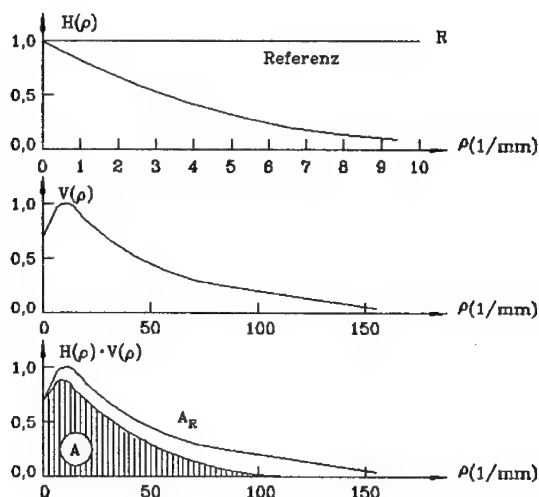


Bild 2.16: Definition der CMT Schärfe.

Die OÜF des optischen Systems $H(\rho)$ ist mit der OÜF des visuellen Systems $V(\rho)$ kaskadiert. Die OÜF vom kompletten System ist dann $H(\rho) \cdot V(\rho)$ im Bild 2.16c. Für die Referenz $R(\rho) = 1$ ist die Kaskadierung mit $V(\rho)$ gleich $R \cdot V(\rho)$. Die CMT-Schärfe ist dann definiert durch

$$\text{CMT-Schärfe} = K + k \cdot \log A/A_R \quad (2.47)$$

A ist die Fläche unter der Funktion $H(\rho) \cdot V(\rho)$ und A_R die Fläche unter der Funktion $R \cdot V(\rho)$.

c. Das Signal-zu-Rauschverhältnis (SNR)-Kriterium

Das SNR-Kriterium ist von C.N. Nelson [14] vorgeschlagen worden. In diesem Kriterium sind die Faktoren zur Beeinflussung der Bildqualität nicht nur die OÜF $H(\rho)$ und $V(\rho)$, sondern auch das Wienerpektrum des gesamten Bildrauschens $W(\rho)$. Zur Definition von SNR ist

$$\frac{S(\rho)}{N(\rho)} = \frac{G(\rho) \cdot H(\rho) \cdot V(\rho)}{\{b(\rho) \cdot (W(\rho) \cdot V^2(\rho) + n)\}^{1/2}} \quad (2.48)$$

wobei $b(\rho)$ die kritische Bandbreite des Rauschens und n der Term für das Rauschen des Auges ist. Das SNR-Kriterium lautet dann:

$$\text{SNR-Kriterium} = \frac{E}{E_R} \quad (2.49)$$

mit

$$E = \int_0^{\infty} \frac{S(\rho)}{N(\rho)} \cdot d\rho$$

und

$$E_R = \int_0^{\infty} \frac{S(\rho)}{N(\rho)} \cdot d\rho \quad , \text{ wenn } H(\rho) = 1 \quad (2.50)$$

Weil alle o.g. Kriterien im Ortsfrequenzbereich erklärt sind, muß das Bild zunächst in diesen Bereich transformiert werden. Diese Transformation ist die Hauptschwierigkeit bei der Berechnung. Um dieses Problem zu vermeiden, sind weitere Kriterien der Bildqualität, die direkt von der Punktantwort oder der Grauwerte des Bildes abhängig sind, vorgeschlagen worden. Der Schärfenindex (1.1) ist ein typisches Kriterium dieser Art.

Die Eigenschaften des zweiten Moments von der Punktantwort als ein Kriterium für die Bildqualität sind seit vielen Jahren untersucht worden. Das zweite Moment der Punktantwort lautet

$$M_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x^2 + y^2) \cdot h(x, y) \cdot dx \cdot dy \quad (2.51)$$

Das zweidimensionale Problem kann auf das eindimensionale Problem reduziert werden, wenn man die Integration in zwei Teile auftrennt. Es wird ein Teil als

$$M_{x,2} = \int_{-\infty}^{\infty} x^2 \cdot l(x) \cdot dx \quad ,$$

definiert, wobei

$$l(x) = \int_{-\infty}^{\infty} h(x, y) \cdot dy$$

die Linienantwort (Englisch: Line Spread Function) längs der x-Achse ist. Das zweite Moment ist dann gegeben durch

$$M_2 = M_{x,2} + M_{y,2} \quad (2.52)$$

mit

$$M_{y,2} = \int_{-\infty}^{\infty} y^2 \cdot l(y) \cdot dy$$

und

$$l(y) = \int_{-\infty}^{\infty} h(x, y) \cdot dx$$

als entsprechende Antwort längs der y-Achse.

3. Automatische Fokussierungsalgorithmen

3.1 Anforderungen an Fokussierungsfunktionen

Die Grundidee der automatischen Fokussierung für die Kamera ist die Suche nach einer Fokussierungsfunktion (FF), die nur von den Grauwerten des durch die Kamera aufgenommenen Bildes abhängig ist. In dem wirksamen Bereich der Kamera $[p'_a, p'_b]$ ist die Fokussierungsfunktion eine unimodulare Kurve (Bild 3.1). In der besten Fokussierungsposition p'_{opt} erreicht die Fokussierungsfunktion einen extremen Wert. In der Nähe von p'_{opt} verändert sich die Fokussierungsfunktion im allgemeinen sehr schnell.

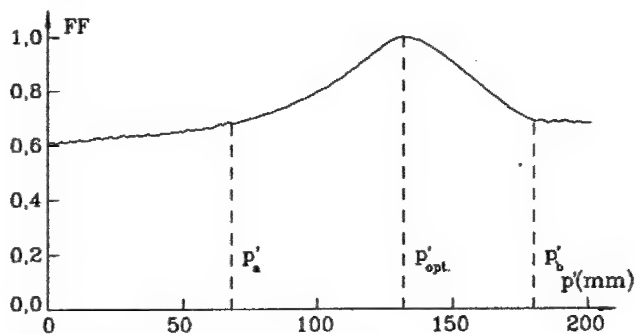


Bild 3.1: Schematische Darstellung einer FF.

Wegen der Wichtigkeit der Fokussierungsfunktion für das automatische Fokussierungssystem werden die folgenden Kriterien zur Auswahl und Beurteilung der Fokussierungsfunktion benutzt.

a. Unimodulare Eigenschaft der Fokussierungsfunktion

Die Fokussierungsfunktion, auf der der Fokussierungsalgorithmus basiert, muß unimodular sein. Das heißt, daß nur ein einziges Maximum oder Minimum in der Fokussierungsfunktion auftreten darf. Diese Bedingung vereinfacht den Algorithmus beträchtlich und vermeidet die potentiellen Fokussierungsfehler von lokalen Extrema.

b. Richtigkeit der optimalen Fokussierungsposition

Das Extrum muß genau in der richtigen Fokussierungsposition der Kamera auftreten.

c. Reproduktionsfähigkeit der FF

Das Extremum der Fokussierungsfunktion muß irgendwann in der richtigen Fokussierungsposition auftreten. Normalerweise führt eine scharfe Spitze des Extrems zur guten Reproduktionsfähigkeit.

d. Die Breite des wirksamen Bereiches $[p'_a, p'_b]$

Die Breite des wirksamen Bereiches $[p'_a, p'_b]$ muß so groß wie möglich sein, damit die Kamera im großen Fokussierungsbereich arbeiten kann.

e. Allgemeine Anwendungsfähigkeit

Die Fokussierungsfunktion muß sich nicht auf einige besondere Arten von Bildern beschränken.

f. Unempfindlichkeit gegenüber anderen Parametern

Wenn sich während des Fokussierungsprozesses andere Parameter verändern, z. B. die Blendenzahl der Kamera oder die Beleuchtungsstärke, sollte der Fokussierungsprozeß nicht durcheinander gebracht werden.

g. Videosignalkompatibilität der FF

Die Fokussierungsfunktion muß auf dem gleichen Videosignal, das als Eingangssignal für die Bildverarbeitung benutzt wird, basieren.

h. Einfachheit der FF

Die Fokussierungsfunktion muß so einfach wie möglich und die Rechenzeit der FF so kurz wie möglich sein.

Die Defokussierung eines optischen Systems kann folgendermaßen modelliert werden: Im Ortsbereich ist ein defokales Bild durch die Faltung des gleichen Bildes im richtigen Fokus mit der Punktantwort des optischen Systems dargestellt. Im Ortsfrequenzbereich ist das Frequenzspektrum eines defokalen Bildes identisch zur Multiplikation des Frequenzspektrums vom gleichen fokalen Bild mit der optischen Übertragungsfunktion. Wenn ein Bild defokussiert ist, nehmen die Leistungsspektren im hohen Frequenzbereich ab, die scharfen Kanten im Bild werden verschliffen und die kleinen Objekte verlieren den Kontrast im Bild. Deshalb kann man die Fokussierungsfunktionen entweder durch die Messung der Leistungsspektren des eingegebenen Bildes im hohen Frequenzbereich oder durch die Messung des Kontrastes oder durch die Messung der Gradientenfunktion der Grauwerte bekommen.

In der vorliegenden Arbeit werden folgende Fokussierungsfunktionen ausgewählt und realisiert.

3.2 Die Varianz der Grauwerte als Fokussierungsfunktion

Die Varianz der Grauwerte von einem Bild $\underline{g} = [g(x,y)]$ mit L Zeilen und R Spalten ist eine Kenngröße, die eine Aussage über den Kontrast im Bild zuläßt. Die Varianz VAR ist definiert:

$$\text{VAR} = \frac{1}{M} \cdot \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} [g(x,y) - m_g]^2, \quad (3.1)$$

wobei m_g der Mittelwert der Grauwerte von \underline{g} ist. Er berechnet sich gemäß

$$m_g = \frac{1}{M} \cdot \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} g(x,y). \quad (3.2)$$

$M = L \cdot R$ ist die Anzahl der Bildpunkte von \underline{g} . Zur einfacheren Berechnung der Grauwerte läßt sich (3.1) durch die Anwendung der binomischen Formel umformen auf:

$$\begin{aligned} \text{VAR} &= \frac{1}{M} \cdot \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} [g(x,y)^2 - 2m_g \cdot g(x,y) + m_g^2], \\ &= \frac{1}{M} \cdot \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} g(x,y)^2 - \frac{2m_g}{M} \cdot \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} g(x,y) + \frac{1}{M} \cdot \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} m_g^2, \\ &= \frac{1}{M} \cdot \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} g(x,y)^2 - m_g^2. \end{aligned} \quad (3.3)$$

Es ist intuitiv, daß die höhere Varianz dem Bild mit der besseren Schärfe entspricht, während die niedrigere Varianz dem unklaren Bild entspricht. Das heißt, daß die Fluktuation der Grauwerte eines fokalen Bildes größer als die von einem defokalen Bild ist. Zur Erklärung der Beziehung zwischen der Defokussierung und der Varianz kann man das Histogramm der relativen Häufigkeit von \underline{g} benutzen.

Entsprechend den fundamentalen Annahmen der Informationstheorie sind Bilder im Rahmen von Verarbeitungsaufgaben als Elemente eines statistischen Ensembles zu betrachten. Zur statistischen Beschreibung von Bildern kann man die klassischen Methoden und Modelle der Theorie des Zufallsprozesses verwenden. Vom Standpunkt dieser Theorie aus ist ein Bild durch seine eindimensionale Wahrscheinlichkeitsdichte statistisch vollständig beschrieben, d. h. durch die Wahrscheinlichkeit, mit der das Bild in ein gegebenes Volumenelement des Signalraumes fällt [15].

Die einfachste und zugleich eine der wichtigsten Charakteristiken von Bildern ist die Wahrscheinlichkeitsverteilung ihrer Grauwerte. Für alle Bildpunkte im digitalen Grauwertbild erhält man eine diskrete Wahrscheinlichkeitsverteilung, das sogenannte Histogramm der Grauwertverteilung. Die Grauwerte $g = g(x,y)$ eines Bildes sind als Realisation einer Zufallsvariablen \tilde{Z} zu interpretieren, wobei hier die Anordnung der Grauwerte im (x,y) -Koordinatensystem keine Rolle spielt. Ein Bild $\underline{G} = [g(x,y)]$ mit der Grauwertmenge $G = \{0,1,2,\dots,255\}$ wird mit einer diskreten Zufallsvariablen \tilde{Z} beschrieben, die die Grauwerte g mit den Wahrscheinlichkeiten p_g , $g = 0,1,\dots, 255$ annimmt. Die Wahrscheinlichkeitsfunktion $g(\tilde{Z})$ ist dann definiert als

$$g(\tilde{Z}) = \begin{cases} p_g & \tilde{Z} = g \\ 0 & \text{für alle übrigen } \tilde{Z} \end{cases}, \quad (3.4)$$

Eine Schätzung der Wahrscheinlichkeitsfunktion $g(\tilde{Z})$ ist das Histogramm der relativen Häufigkeiten

$$g(\tilde{Z}) = \begin{cases} p(g) = \frac{a_g}{M} & \tilde{Z} = g \\ 0 & \text{für alle übrigen } \tilde{Z} \end{cases}. \quad (3.5)$$

Dabei ist a_g die Häufigkeit des Auftretens des Grauwertes g im Bild. Mit dem Zufallsprozeß ist der Mittelwert m_g als der Erwartungswert der Zufallsgröße \tilde{Z} zu berechnen, also

$$m_g = E(\tilde{Z}) = \sum_{g=0}^{255} g \cdot p_g, \quad (3.6)$$

und die Varianz der Grauwerte von \tilde{Z} ist

$$VAR = E[(\tilde{Z} - E(\tilde{Z}))^2] = \sum_{g=0}^{255} [g - E(\tilde{Z})]^2 \cdot p_g. \quad (3.7)$$

Eine Schätzung von $E(\tilde{Z})$ und $E[(\tilde{Z} - E(\tilde{Z}))^2]$ ist durch das Histogramm realisierbar, nämlich

$$m_g = E(\tilde{Z}) = \sum_{g=0}^{255} g \cdot p(g), \quad (3.8)$$

$$VAR = E[(\tilde{Z} - E(\tilde{Z}))^2] = \sum_{g=0}^{255} (g - m_g)^2 \cdot p(g). \quad (3.9)$$

Bild 3.2 zeigt das Profil einer idealen Kante längs der x-Achse und sein Histogramm. In Bild 3.2a ist die Kamera richtig fokussiert, im Bild tritt ein Grauwertsprung auf. Das Histogramm besteht aus zwei δ -Funktionen. Im Fall der besten Fokussierung ist das Histogramm

$$p(g) = 0,5 \cdot [\delta(g-a) + \delta(g-b)].$$

Der Mittelwert und die Varianz sind

$$m_{g1} = \int_0^L p(g) \cdot g \cdot dg = 0,5 \cdot (a+b),$$

$$VAR_1 = \int_0^L [g - 0,5 \cdot (a+b)]^2 \cdot p(g) \cdot dg = 0,25 \cdot (b-a)^2.$$

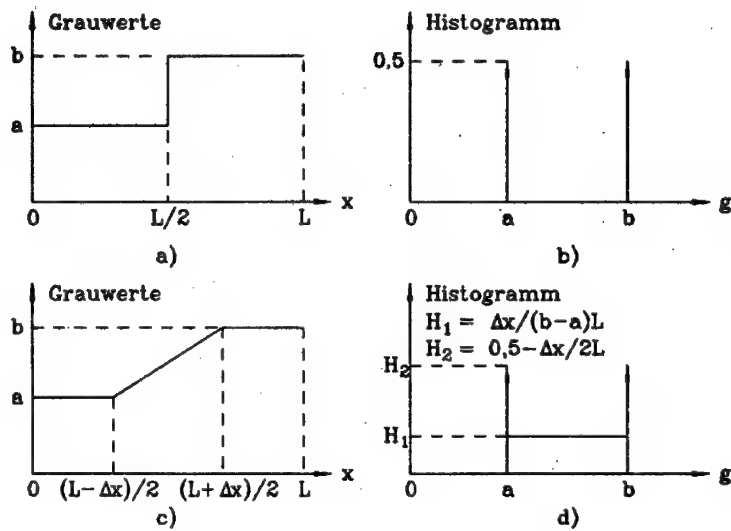


Bild 3.2: Die Grauwerte und das Histogramm des Profils einer Kante.

In Bild 3.2c hat die Kante durch Defokussierung zum Verschleifen des Grauwertesprungs geführt. Das Histogramm läßt sich durch folgende Gleichung darstellen

$$p(g) = \begin{cases} 0 & g < a \\ (0.5 - \frac{\Delta x}{2L}) \cdot \delta(g-a) & g = a \\ \frac{\Delta x}{L \cdot (b-a)} & a < g < b \\ (0.5 - \frac{\Delta x}{2L}) \cdot \delta(g-b) & g = b \\ 0 & g > b \end{cases}$$

Man kann den Mittelwert und die Varianz von dieser defokalen Kante durch

$$m_{g2} = 0.5 \cdot (a+b) ,$$

$$VAR_2 = 0.25 \cdot (b-a)^2 - \frac{\Delta x}{6L} \cdot (b-a)^2$$

bekommen. Es ist klar, daß die Varianz der defokalen Kante VAR_2 kleiner als die der fokalen Kante VAR_1 ist. Normalerweise, wenn sich der Fokussierungszu-

stand verschlechtert, wird Δx sich vergrößern. An VAR_2 kann man dann bemerken, daß die Varianz sich verkleinert.

Bei einem Bild mit reichem Kontrast ist vor allem das Histogramm $p(g)$ für Werte von g im breiten Bereich besetzt. Das Bild hat eine große Varianz. Im Gegensatz beim Bild mit wenig Kontrast konzentriert sich das Histogramm $p(g)$ in der Nähe des Mittelwertes. Das Bild besitzt eine kleine Varianz.

3.3 Die Entropie der Grauwerte als Fokussierungsfunktion

Die Entropie ist ein Maß des mittleren Informationsgehalts eines Bildes. Wenn das Bild als eine empfangene Nachricht (message), die die Information über das Objekt liefert, betrachtet wird, kann man die Bildqualität nach der Informationstheorie von Shannon auswerten. Die Grundidee läßt sich folgendermaßen charakterisieren:

Jede Nachrichtenquelle (Geber) sendet energetische Signale über eine Übertragungskette an den Empfänger. Der Empfänger ist im weitesten Sinne nicht ein physikalisches Gerät, sondern ein Sinnesorgan des Beobachters. Jeder Geber ist imstande, eine bestimmte Anzahl von Nachrichten zu übertragen, er besitzt einen mittleren Informationsgehalt. Das gleiche gilt für den Empfänger, der ebenfalls nur eine bestimmte Anzahl von Nachrichten aufnehmen kann (mittlerer Informationsgehalt des Empfängers). Der mittlere Informationsgehalt bezieht sich also nicht auf die etwa zu einem bestimmten Zeitpunkt vom Geber ausgesandten oder vom Empfänger aufgenommenen Nachrichten, sondern auf die maximal möglichen Nachrichten. Der die Verbindung zwischen Geber und Empfänger herstellende Übertragungskanal kann nur eine bestimmte Anzahl von Nachrichten weiterleiten; er besitzt eine Übertragungskapazität.

Bei der rechnerischen Erfassung der Nachrichtenübertragung kann man die Entropie benutzen. Voraussetzung ist dabei allerdings, daß die Nachrichten eine stochastische Menge bilden. Dazu gehört, daß jede einzelne Nachricht durch ihre Wahrscheinlichkeit gekennzeichnet ist. Der Informationsgehalt einer Einzelnachricht ist um so größer, je geringer seine Häufigkeit ist. Diese Forderung erscheint durchaus plausibel, wenn man bedenkt, daß der Wert einer Nachricht mit ihrer Seltenheit steigt. Der mittlere Informationsgehalt des Senders (oder auch des Empfängers) ist durch die durchschnittliche Zahl der Nachrich-

teneinheiten bestimmt. Man kann für den mittleren Informationsgehalt die Entropie ENT definieren, die durch die folgende Formel gegeben ist

$$ENT = - \sum_{i=1}^n p_i \cdot \log_2(p_i) \quad (3.10)$$

In dieser Formel bedeutet p_i die Wahrscheinlichkeit für die i -te Nachricht. Da jede Nachricht durch eine bestimmte Anzahl von Entscheidungen zwischen "ja" und "nein" charakterisiert ist, kann man die Ja-Nein-Entscheidung als Nachrichteneinheit festlegen. Man spricht dabei von Bits und rechnet in Zweier-Logarithmen.

Für die Bildverarbeitung steht nach der Digitalisierung jedem Bildpunkt des Bildes eine Menge $\{0,1,2,\dots,255\}$ von insgesamt 256 Graustufen zur Verfügung, d.h. ein Bild hat 256 Nachrichten. Es sei angenommen, daß es möglich ist, für den Grauwert g eine zugehörige Wahrscheinlichkeit p_g anzugeben. Der Grauwert g und die zugehörige Wahrscheinlichkeit p_g bilden ein Ensemble. Der mittlere Informationsgehalt ist dann durch die Entropie gegeben

$$ENT = - \sum_{g=0}^{255} p_g \cdot \log_2(p_g) \quad (3.11)$$

Für ein homogenes Bild ist die Entropie $ENT = 0$, weil im ganzen Bild nur ein Grauwert auftritt.

$$p_g = \begin{cases} 1 & g = \text{ein bestimmter Wert} \\ 0 & \text{sonst.} \end{cases}$$

Im Gegensatz dazu erreicht ein Bild, in dem alle Grauwerte mit derselben relativen Wahrscheinlichkeit $p_g = 1/256$ auftreten, die maximale Entropie, nämlich $ENT = 8$. Für die meisten Situationen konzentriert sich das Histogramm vom defokalen Bild auf einige Grauwerte in der Nähe vom Mittelwert (Bild 3.3a).

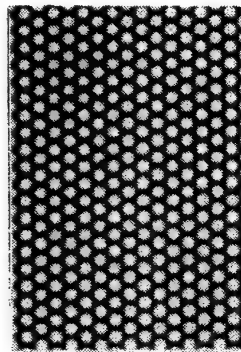
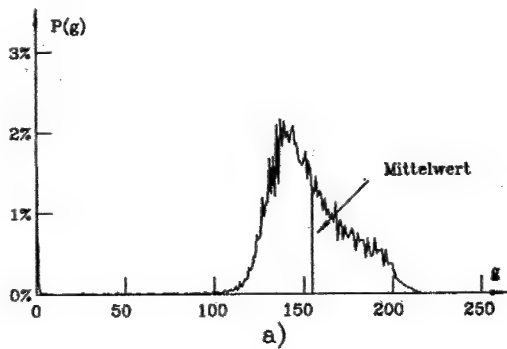


Bild 3.3: Histogramm eines defokalen Bildes.

Deshalb hat ein defokales Bild einen niedrigen mittleren Informationsgehalt, nämlich eine kleine Entropie. Im Gegensatz dazu besitzt ein fokales Bild eine große Entropie.

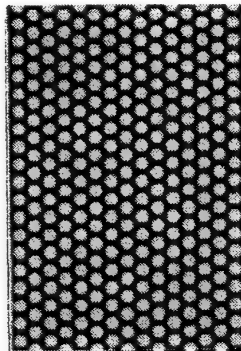
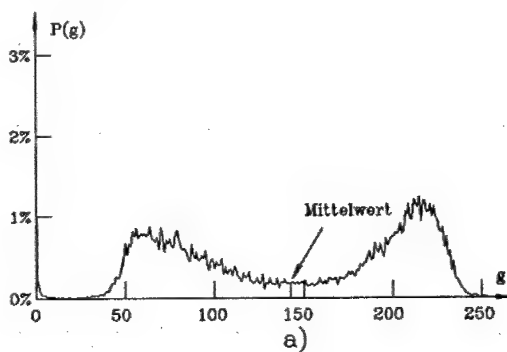


Bild 3.4: Histogramm eines fokalen Bildes.

Wie Bild 3.4a zeigt, verteilt sich das Histogramm über breite Intervalle in der Grauwertmenge. In der besten Fokussierungsposition des optischen Systems erreicht die Entropie das Maximum.

3.4 Die SMD der Grauwerte als Fokussierungsfunktion

SMD ist eine Abkürzung für "Sum Modulus Difference", die ein Maß für den absoluten Gradienten des Bildes darstellt. In einem zweidimensionalen Koordinatensystem ist der Gradient eine stetige Funktion zweier Variablen

$$\begin{bmatrix} \frac{\partial g(x,y)}{\partial x} & , & \frac{\partial g(x,y)}{\partial y} \end{bmatrix}^T \quad (3.12)$$

Der Betrag des Gradienten ist

$$|\nabla g(x,y)| = \sqrt{\left[\frac{\partial g}{\partial x}\right]^2 + \left[\frac{\partial g}{\partial y}\right]^2} \quad (3.13)$$

und die Richtung des Gradienten berechnet sich aus

$$\theta(\nabla g(x,y)) = \arctan \left[\frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}} \right] \quad (3.14)$$

Durch den Gradientenoperator werden kleine Strukturen im Bild hervorgehoben. Bild 3.5 zeigt, daß beispielsweise der Gradientenoperator für die Kantendetektion geeignet ist. Der Gradient hat an einer Kante ein Maximum. Im homogenen Bereich ist der Gradient Null.

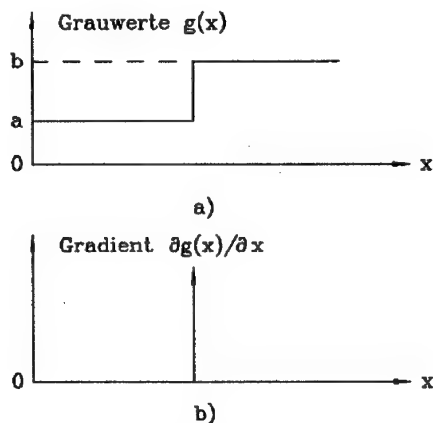


Bild 3.5: Eindimensionale Grauwertkante und ihr Gradient.

Im Diskreten kann der Gradient nur approximiert werden. Die Herstellung eines Gradienten erfordert in jedem Bildpunkt die Differenzierung in zwei orthogonale Richtungen der Ortskoordinaten. Weil der Grauwert diskret ist, müssen die Differentialquotienten durch endliche Differenzen ersetzt werden

$$\begin{cases} \frac{\partial g(x,y)}{\partial x} = \frac{g(x,y) - g(x-1,y)}{x - (x-1)} = g(x,y) - g(x-1,y) \\ \frac{\partial g(x,y)}{\partial y} = \frac{g(x,y) - g(x,y-1)}{y - (y-1)} = g(x,y) - g(x,y-1) \end{cases} \quad (3.15)$$

$\partial g(x,y)/\partial x$ und $\partial g(x,y)/\partial y$ sprechen auf horizontal und vertikal verlaufende Grauwertkanten entsprechend an.

Die Fokussierungsfunktion SMD hat drei Formen. Die erste Form

$$SMD1 = \sum_{y=0}^{R-1} \sum_{x=0}^{L-1} \left| \frac{\partial g(x,y)}{\partial x} \right| = \sum_{y=0}^{R-1} \sum_{x=0}^{L-1} |g(x,y) - g(x-1,y)| \quad (3.16)$$

extrahiert den Kantengradienten entlang der x-Richtung. Gleichfalls extrahiert die zweite Form

$$SMD2 = \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} \left| \frac{\partial g(x,y)}{\partial y} \right| = \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} |g(x,y) - g(x,y-1)| \quad (3.17)$$

den Kantengradienten entlang der y-Richtung. Die Fokussierungsfunktionen SMD1 und SMD2 haben einen Nachteil. Für einige Bilder gibt es Kanten nur entlang der x-Richtung oder y-Richtung (Bild 3.6). In dieser Situation sind die Gradienten entlang der y-Richtung oder der x-Richtung gleich 0.

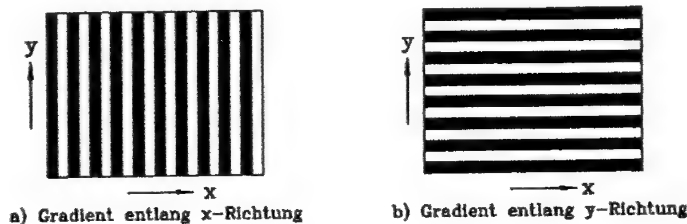


Bild 3.6: Zwei spezielle Kantenrichtungen.

Zur Vermeidung dieses Problems kann der Betrag des Gradienten als Fokussierungsfunktion benutzt werden

$$\begin{aligned} \text{SMD3} &= \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} \sqrt{\left[\frac{\partial g(x,y)}{\partial x}\right]^2 + \left[\frac{\partial g(x,y)}{\partial y}\right]^2} \quad (3.18) \\ &= \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} \sqrt{(g(x,y)-g(x-1,y))^2 + (g(x,y)-g(x,y-1))^2} \end{aligned}$$

Ideale Ableitungsoperatoren ergeben sich im zweidimensionalen Koordinatensystem aus folgenden Operatoridentitäten [16]:

$$\begin{cases} \frac{\partial}{\partial x} g(x,y) \longleftrightarrow j u \cdot G(u,v) \\ \frac{\partial}{\partial y} g(x,y) \longleftrightarrow j v \cdot G(u,v) \end{cases} \quad (3.19)$$

Im Ortsfrequenzbereich bedeutet die Differentiation eine Multiplikation des Frequenzspektrums mit der Ortsfrequenz, d. h. es erfolgt eine Anhebung der hohen Frequenzspektren. Es wird deutlich, daß für ein defokales Bild die Leistungsspektren im hohen Frequenzbereich wegen der Defokussierung gefiltert sind, deswegen sind die Grauwertkanten nicht mehr so deutlich ausgeprägt wie die im fokalen Bild.

3.5 Die Leistungsspektren des Bildes als Fokussierungsfunktion

Die Leistungsspektren im hohen Ortsfrequenzbereich vom defokalen Bild werden unterdrückt. Man kann deshalb die Leistungsspektren direkt benutzen, um die Fokussierungsfunktion zu berechnen. Dazu muß eine zweidimensionale diskrete Fouriertransformation (DFT) durchgeführt werden. Man erhält so die diskrete Fouriertransformierte

$$G(u,v) = \frac{1}{L \cdot R} \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} g(x,y) \cdot \exp\left[-2\pi \cdot j \cdot \left(\frac{x \cdot u}{L} + \frac{y \cdot v}{R}\right)\right] \quad (3.20)$$

Das Leistungsspektrum in der Ortsfrequenz (u,v) ist

$$G_1(u,v) = |G(u,v)|^2 \quad (3.21)$$

Die Fokussierungsfunktion kann durch die Summation der Leistungsspektren im hohen Frequenzbereich erzielt werden. Es ist

$$LS = \sum_{u,v \in \mathbb{H}} G_1(u,v) \quad . \quad (3.22)$$

\mathbb{H} ist die Fläche des hohen Frequenzbereiches. In der vorliegenden Arbeit besteht sie aus allen Ortsfrequenzen außerhalb des Ursprungs $(u,v) = (0,0)$.

Die zweidimensionale diskrete Fouriertransformation ist ein zeitraubendes Verfahren. Wenn man eine zweidimensionale DFT eines $N \times N$ Bildes direkt nach (3.20) berechnet, erfordert jeder Bildpunkt des transformierten Bildes, das ja durch die Multiplikation der entsprechenden Basisfunktion $\exp(-2\pi \cdot j \cdot (x \cdot u/N + y \cdot v/N))$ mit allen Bildpunkten des Bildes entsteht, N^2 komplexe Multiplikationen und $N^2 - 1$ komplexe Additionen. Die Durchführung der gesamten DFT erfordert daher insgesamt N^4 komplexe Multiplikationen und $N^2 \cdot (N^2 - 1)$ komplexe Additionen. Mit der schnellen Fouriertransformation (FFT) kann die Rechenzeit enorm reduziert werden. Dazu wird die zweidimensionale DFT in zwei Schritten berechnet: es erfolgt nämlich die Zerlegung in eine Aufeinanderfolge von eindimensionalen DFT. Wegen der Separierbarkeit der Basisfunktion kann die Gleichung (3.20) entsprechend umgeformt werden

$$G(u,v) = \frac{1}{N} \cdot \sum_{x=0}^{N-1} \left[\frac{1}{N} \cdot \sum_{y=0}^{N-1} g(x,y) \cdot \exp(-2\pi \cdot j \cdot y \cdot v/N) \right] \cdot \exp(-2\pi \cdot j \cdot x \cdot u/N) \quad .$$

Die innere Summe stellt eine eindimensionale DFT der N Spalten dar, die äußere die anschließende Transformation der N Zeilen. Die 2D-DFT läßt sich damit als Aufeinanderfolge zweier 1D-DFT ausführen, N Zeilentransformationen folgen N Spaltentransformationen auf dem $N \times N$ Bild:

$$\text{Spaltentransformation } G(x,v) = \frac{1}{N} \cdot \sum_{y=0}^{N-1} g(x,y) \cdot \exp(-\frac{2\pi \cdot j \cdot y \cdot v}{N}) \quad ,$$

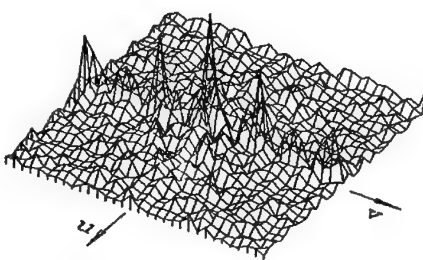
$$\text{Zeilentransformation } G(u,v) = \frac{1}{N} \cdot \sum_{x=0}^{N-1} G(x,v) \cdot \exp(-\frac{2\pi \cdot j \cdot x \cdot u}{N}) \quad .$$

Für jede der Teiltransformationen kann der 1D-FFT Algorithmus eingesetzt werden. Da zwei Transformationen hintereinander ausgeführt werden, hat sich pro Bildpunkt gerechnet, der Aufwand gegenüber der 1D-DFT nur verdoppelt.

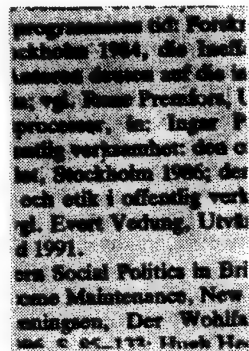
Für die qualitative Beurteilung der in einem Bild vorhandenen Ortsfrequenzen wird für das Leistungsspektrum $|G(u,v)|^2$ häufig eine dreidimensionale Darstellung angestrebt. Sollen in der dreidimensionalen Darstellung Spektralkomponenten mit kleinem Betrag stärker betont werden als solche mit großem Betrag, so kann die Funktion $G'(u,v)$ einer dreidimensionalen Darstellung von $|G(u,v)|^2$ berechnet werden gemäß

$$G'(u,v) = c \cdot \log(1 + |G(u,v)|^2) \quad , \quad (3.23)$$

wobei c ein Normierungsfaktor ist. Rechts in Bild 3.7 ist ein fokales Bild gezeigt und links die dreidimensionale Darstellung des Leistungsspektrums.



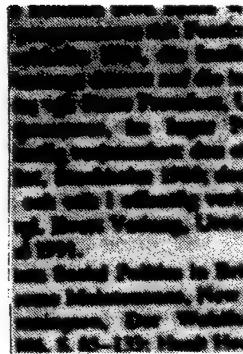
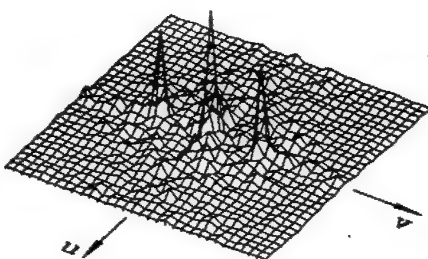
a)



b)

Bild 3.7: Leistungsspektrum vom fokalen Bild.

In diesem Fall sind die Leistungsspektren im hohen Frequenzbereich nicht unterdrückt. Im Gegensatz dazu zeigt das Bild 3.8 ein defokales Bild und die entsprechend dreidimensionale Darstellung des Spektrums. In dieser Situation sind die Leistungsspektren im hohen Frequenzbereich verschwunden bzw. stark unterdrückt.



Bestmögliche Kopie

Bild 3.8: a) Dreidimensionale Darstellung des Leistungsspektrums vom defokalen Bild. b)

3.6 Die Signalleistung der Grauwerte als Fokussierungsfunktion

Die Signalleistung des Bildes ist durch folgende Gleichung definiert:

$$SL = \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} g(x,y)^2 \quad (3.24)$$

SL ist proportional zur Energie im Bild. Für das fokale Bild sind die Energien im hohen Frequenzbereich nicht gefiltert, die SL wird ihr Maximum erreichen.

Wenn die OÜF annähernd als Gaußverteilung dargestellt ist, kann man zeigen, daß die SL das Maximum im besten Fokussierungszustand des Bildes erreicht. Normalerweise erfüllen die Bilder das Parseval-Theorem für die zweidimensionale DFT

$$\sum_{x=0}^{L-1} \sum_{y=0}^{R-1} g(x,y)^2 = L \cdot R \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} G(u,v) \cdot G^*(u,v) \quad (3.25)$$

Im Ortsfrequenzbereich ergibt sich

$$G(u,v) = F(u,v) \cdot H(u,v) = F(u,v) \cdot \exp\left[-\frac{(u^2+v^2) \cdot \sigma^2}{2}\right] \quad (3.36)$$

$F(u,v)$ ist die Fouriertransformierte vom Eingangsbild oder vom fokalen Bild. Für eine bestimmte Szene ist $F(u,v)$ unverändert, während sich der Fokussierungszustand vom optischen System verändert. Weiter ist

$$\begin{aligned} G(u,v) \cdot G^*(u,v) &= F(u,v) \cdot F^*(u,v) \cdot \exp(-(u^2+v^2) \cdot \sigma^2) \\ &= |F(u,v)|^2 \cdot \exp(-(u^2+v^2) \cdot \sigma^2) \end{aligned} \quad (3.27)$$

Setzt man (3.27) in (3.25) ein, so ergibt sich

$$SL = L \cdot R \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |F(u,v)|^2 \cdot \exp(-(u^2+v^2) \cdot \sigma^2)$$

Weil $F(u,v)$ für eine unveränderte Szene bestimmt und $|F(u,v)|^2 > 0$ ist, ist SL nur von dem Ausdehnungsparameter σ abhängig.

Wenn SL ein Extremum erreicht, muß

$$\frac{d(SL)}{d\sigma} = 0$$

gelten, wobei

$$\frac{d(SL)}{d\sigma} = -2\sigma \cdot R \cdot L \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |F(u,v)|^2 \cdot (u^2+v^2) \cdot \exp(-(u^2+v^2) \cdot \sigma^2) \quad (3.28)$$

ist.

Wenn das Bild homogen ist, also alle Grauwerte des Bildes identisch sind, haben alle Leistungsspektren bis auf das Leistungsspektrum $|F(0,0)|^2$ den Wert 0. Deshalb ist

$$2R \cdot L \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |F(u,v)|^2 \cdot (u^2+v^2) \cdot \exp(-(u^2+v^2) \cdot \sigma^2)$$

immer Null, d.h., die Fokussierungsfunktion SL ist unabhängig von dem Ausdehnungsparameter σ . In diesem Fall hat SL die Fähigkeit zur Fokussierung verloren. Im Gegensatz dazu besitzt ein inhomogenes Bild Leistungsspektren im hohen Frequenzbereich. In diesem Fall ist

$$2R \cdot L \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |F(u,v)|^2 \cdot (u^2+v^2) \cdot \exp(-(u^2+v^2) \cdot \sigma^2) > 0.$$

Das bedeutet, nur wenn σ gleich 0 ist, erreicht SL das Extremum. Ein von Null identisches σ bedeutet, daß das optische System im besten Fokussierungszustand ist. Außerdem ist

$$\frac{d^2(SL)}{d\sigma^2} = - 2R \cdot L \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |F(u,v)|^2 \cdot (u^2+v^2) \cdot [1-2(u^2+v^2) \cdot \sigma^2] \cdot \exp[-(u^2+v^2) \cdot \sigma^2]$$

Im Fall $\sigma = 0$ ergibt sich

$$\frac{d^2(SL)}{d\sigma^2} = - 2R \cdot L \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |F(u,v)|^2 \cdot (u^2+v^2) < 0$$

Das bedeutet, daß SL einen maximalen Wert im besten Fokussierungszustand des optischen Systems erreicht.

In der Praxis ist die SL als Fokussierungsfunktion nicht günstig, weil die Kurve von SL zu flach ist (Bild 3.9).

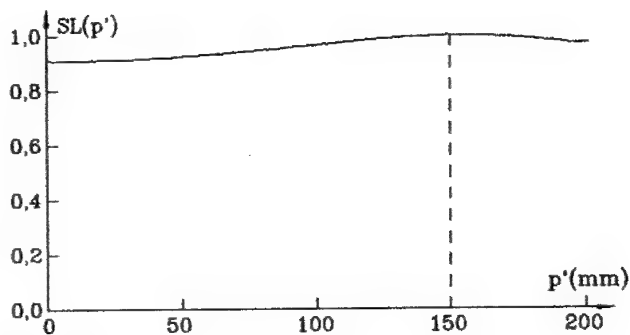


Bild 3.9: SL als Fokussierungsfunktion.

Um diesen Nachteil zu vermeiden, wird eine veränderte SL als Fokussierungsfunktion benutzt. Es ist dies

$$SL' = \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} g'(x,y)^2, \quad (3.29)$$

wobei $g'(x,y)$ durch folgende Gleichung definiert ist:

$$g'(x,y) = \begin{cases} g(x,y) & \text{wenn } g(x,y) < g^* \\ 0 & \text{sonst} \end{cases}, \quad (3.30)$$

g^* ist eine Konstante, die durch das Experiment bestimmt ist. Bild 3.10 zeigt den Verlauf von SL' .

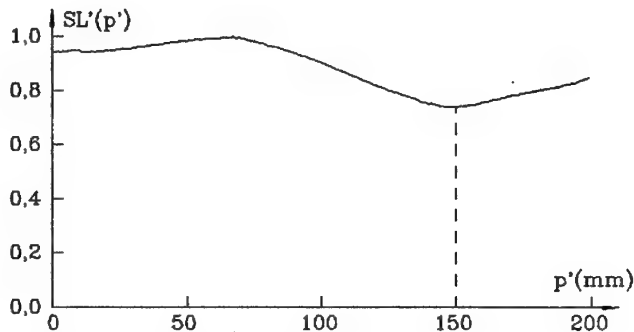


Bild 3.10: SL' als Fokussierungsfunktion.

Beim Vergleich zwischen Bild 3.9 und Bild 3.10 bemerkt man, daß die besten Fokussierungspositionen beider Kurven keinen Unterschied haben, aber die SL' erreicht nicht den maximalen Wert sondern einen minimalen Wert.

3.7 "Thresholded Video Signal Pixel Count" als Fokussierungsfunktion

Die Fokussierungsfunktion "Thresh..." (TV) ist die Zahl der Bildpunkte im Bild, deren Grauwerte über oder unter einem bestimmten Schwellenwert liegen. TV berechnet sich gemäß:

$$TV = \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} S[g(x,y), \mu, c], \quad (3.31)$$

wobei $S[g(x,y),\mu,c]$ durch die Gleichung

$$S[g(x,y),\mu,c] = \begin{cases} 1 & \text{wenn } |g(x,y) - \mu| \geq c \\ 0 & \text{sonst} \end{cases} \quad (3.32)$$

definiert ist. μ ist eine Variable, die vom Bild abhängig ist. In der vorliegenden Arbeit wird μ als der Mittelwert des Bildes gewählt, c ist eine Konstante, die normalerweise durch das Experiment bestimmt ist. Ein zu großes oder zu kleines c ergibt eine schlechte Fokussierungsfunktion. Das Fokussierungsprinzip der TV kann durch die Defokussierungswirkung auf eine eindimensionale Grauwertkante erklärt werden.

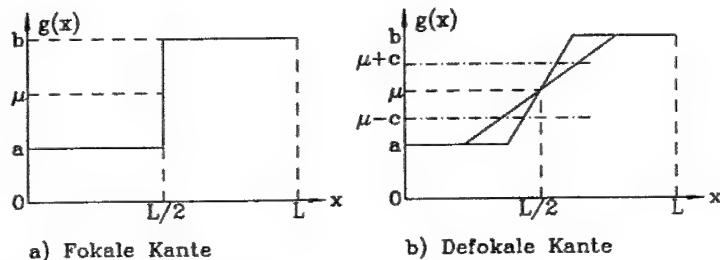


Bild 3.11: Eindimensionale Kante.

Im Fall der richtigen Fokussierung ergibt sich eine ideale Grauwertkante (Bild 3.11a). Der Mittelwert der Kante ist $(a+b)/2$. Wenn die Konstante c so gewählt wird, daß $c < (a-b)/2$ erfüllt ist, dann erreicht die Fokussierungsfunktion $TV = L$. Aber wenn die Kante defokussiert ist, wie in Bild 3.11b gezeigt, sind die Grauwerte verschliffen. Die Grauwerte von vielen Bildpunkten befinden sich innerhalb des Intervalls $[\mu-c, \mu+c]$. Die Defokussierungsfunktion ist in dieser Situation kleiner als L . Je schlechter der Fokussierungszustand ist, desto schiefer wird die Kante, und es befinden sich immer mehr Bildpunkte im Intervall $[\mu-c, \mu+c]$.

Wenn die Häufigkeit des Auftretens der Grauwerte $p(g)$ eine Funktion von g ist, ist der Wert TV proportional zur Fläche der gestrichelten Gebiete in Bild 3.12 und Bild 3.13, also

$$TV = L \cdot R \cdot \left[\int_0^{\mu-c} p(g) \cdot dg + \int_{\mu+c}^{255} p(g) \cdot dg \right] \quad (3.33)$$

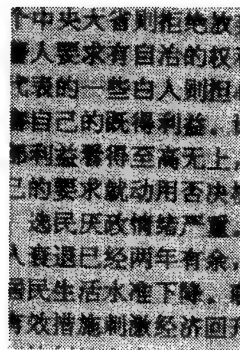
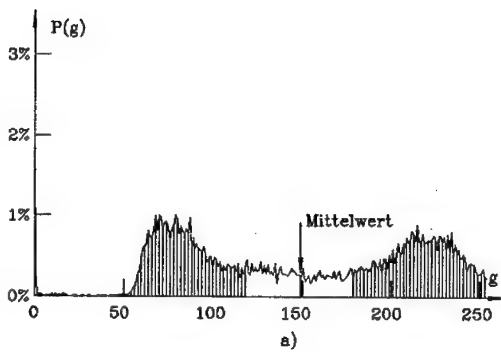


Bild 3.12: Histogramm des fokalen Bildes.

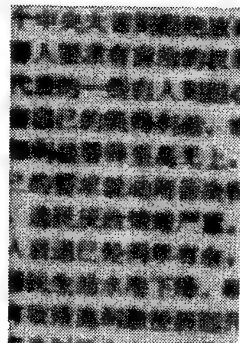
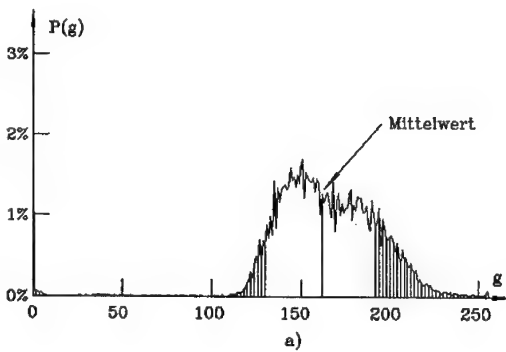


Bild 3.13: Histogramm des defokalen Bildes.

Für ein defokales Bild konzentriert sich das Histogramm in der Nähe vom Mittelwert der Grauwerte. Die Fläche der gestrichelten Gebiete (Bild 3.13) ist dann kleiner als die bei einem fokalen Bild.

Es ist klar, daß die Auswahl der Konstante c sehr wichtig für die Fokussierungsfunktion TV ist. Wenn c zu groß ist, wird $TV = 0$. Im Gegensatz eines zu kleinen c verläuft die Kurve der Fokussierungsfunktion zu flach.

Die Konstante c kann aus dem Histogramm des fokalen Bildes ermittelt werden. Normalerweise wird sich beim fokalen Bild ein bimodales Histogramm ergeben. Der Mittelwert befindet sich zwischen den beiden Maxima. Die Konstante c muß dann kleiner als $|\mu - g_{\max 1}|$ und $|\mu - g_{\max 2}|$ sein (Bild 3.12).

3.8 Der Laplace-Operator der Grauwerte als Fokussierungsfunktion

Wie gezeigt wurde, liegt der Unterschied zwischen dem fokalen und dem defokalen Bild bei hohen Frequenzen. Das zweite Moment der Fourierspektren ist ein Maß für die hohen Frequenzkomponenten des Bildes

$$L_0 = k \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |(u^2 + v^2) \cdot G(u, v)|^2 \quad (3.34)$$

In der Fokussierungsfunktion L_0 sind die hohen Frequenzkomponenten stark hervorgehoben. Wenn die Konstante k als

$$k = (4 \cdot \pi^2)^2 \cdot L \cdot R$$

gewählt wird, ergibt sich

$$\begin{aligned} L_0 &= L \cdot R \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |(4\lambda^2 \cdot u^2 + 4\pi^2 \cdot v^2) \cdot G(u, v)|^2, \\ &= L \cdot R \cdot \sum_{u=0}^{L-1} \sum_{v=0}^{R-1} |(-4\pi^2 \cdot u^2 - 4\pi^2 \cdot v^2) \cdot G(u, v)|^2. \end{aligned} \quad (3.35)$$

Im Zweidimensionalen gilt folgende Operatoridentität:

$$\nabla^2 g(x, y) \longrightarrow (-4\pi^2 \cdot u^2 - 4\pi^2 \cdot v^2) \cdot G(u, v) \quad (3.36)$$

Wegen des Parseval-Theorems bekommt man folglich

$$L0 = \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} \left[\nabla^2 g(x,y) \right]^2 \quad (3.37)$$

Im Zweidimensionalen ist der Laplace-Operator

$$\nabla^2 g(x,y) = \frac{\partial^2 g(x,y)}{\partial x^2} + \frac{\partial^2 g(x,y)}{\partial y^2} \quad (3.38)$$

Die Approximation der zweiten Ableitung durch Differenzen ergibt

$$\frac{\partial^2 g(x,y)}{\partial x^2} = g(x+1,y) - 2g(x,y) + g(x-1,y) \quad ,$$

$$\frac{\partial^2 g(x,y)}{\partial y^2} = g(x,y+1) - 2g(x,y) + g(x,y-1) \quad .$$

Setzt man $\nabla^2 g(x,y)$ in (3.37) ein, so folgt

$$L0 = \sum_{x=0}^{L-1} \sum_{y=0}^{R-1} \left[g(x+1,y) + g(x-1,y) + g(x,y+1) + g(x,y-1) - 4g(x,y) \right]^2 \quad (3.39)$$

Im Bild entfernt der Laplace-Operator nicht nur die konstanten Grauwertbereiche, sondern auch jede konstante Grauwertsteigung. Es ist zu beachten, daß der Laplace-Operator, angewendet auf eine beliebig geneigte Ebene, den Wert 0 ergibt, obwohl die Gradienten dieser Fläche nicht den Wert 0 haben.

3.9 Das Differential der Varianz als Fokussierungsfunktion

Die Varianz der Grauwerte VAR als Funktion der Fokussierungsposition p' hat ein sehr gutes unimodulares Verhalten im wirksamen Arbeitsbereich der Kamera. Als unimodulare Funktion erreicht VAR(p') das Extremum in der Position, wo

$$\frac{dVAR(p')}{dp'} = 0$$

ist (Bild 3.14).

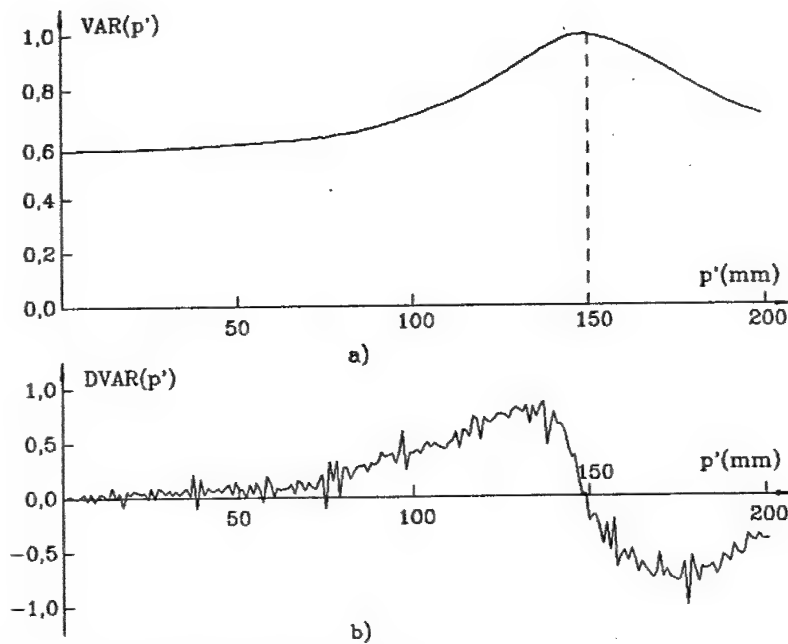


Bild 3.14: Varianz und ihr Differential.

Als die ideale unimodulare Funktion nimmt VAR auf der linken Seite der besten Fokussierungsposition monoton zu und auf der rechten Seite entsprechend ab. Deshalb schneidet das Differential der Varianz in der besten Fokussierungsposition die p' -Achse. Um die beste Fokussierungsposition zu finden, braucht man also nur den Nulldurchgangspunkt in der p' -Achse zu bestimmen. Als Fokussierungsfunktion lautet das Differential der Varianz

$$\text{DVAR}(p') = \frac{d\text{VAR}(p')}{dp'} \approx \text{VAR}(p') - \text{VAR}(p'-1) \quad . \quad (3.40)$$

Aber fast alle Bilder sind verrauscht, besonders im unwirksamen Arbeitsbereich der Kamera. Wie Bild 3.14 zeigt, geht $\text{DVAR}(p')$ mehrmals durch die Abszisse.

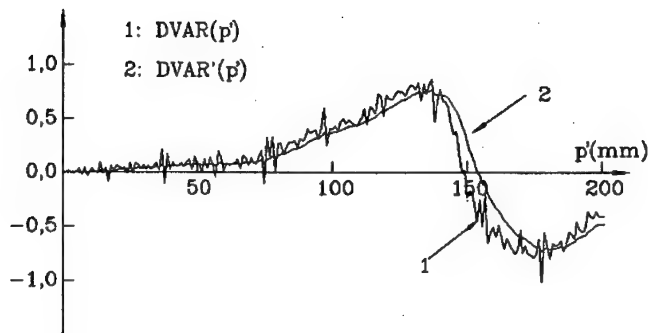


Bild 3.15: Auswirkung des Filters.

Um dieses Problem zu lösen, wird ein Tiefpaßfilter benutzt. Das einfachste Tiefpaßfilter hat die Form

$$DVAR'(p') = a \cdot DVAR(p') + b \cdot DVAR(p'-1) \quad , \quad (3.41)$$

wobei a , b die Konstanten sind, die die Forderung $a + b = 1$ erfüllen. Die Wahl von a , b ist sehr wichtig für das Filterergebnis. Wenn a zu klein ist, ist das Rauschen nicht gut herausgefiltert. Bei einem zu großen a wird die Funktion $DVAR'(p')$ mit großer Verzögerung auftreten (Bild 3.15). Diese Verzögerung kann allerdings durch Experimente bestimmt und im Rechenprogramm kompensiert werden.

4. Beurteilung der verschiedenen Fokussierungsfunktionen durch Experimente

Zur Beurteilung der Vor- und Nachteile der verschiedenen Fokussierungsfunktionen werden Experimente durchgeführt. Die Ergebnisse der Experimente sind durch die Kurven der FF dargestellt. In den Experimenten wird die Position des zu fokussierenden Objektes nicht verändert, nur die Kamera bewegt sich. Ein computergesteuerter Schrittmotor wird benutzt zur Kontrolle der Bewegung der Kamera. Das Objekt befindet sich 225 mm weit von der Anfangsposition der Kamera. Die Kamera tastet die Szene von 0 mm bis zu 200 mm ab. Jeder Abtastschritt ist 1 mm, d.h. jede Kurve der FF hat 201 Abtastpunkte. Um so viele Daten zu behandeln, ist ein automatisches Meßprogramm zu entwickeln. Nach der Abtastung der Szene wird die Fokussierungsfunktion berechnet.

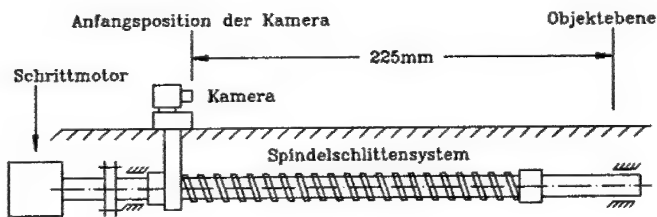


Bild 4.1: Skizze der experimentellen Vorrichtung.

Um die grafische Darstellung der Meßergebnisse zu erleichtern, werden alle Fokussierungsfunktionen durch den jeweiligen Maximalwert normiert. Deshalb ist das Maximum der Fokussierungsfunktionen immer 1. Es ist angenommen, daß die zu fokussierenden Objekte eine ebene Oberfläche besitzen oder deren Schnittdicke kleiner als die Schärfentiefe der Kamera ist.

Um den Einfluß des Rauschens zu unterdrücken, sollten günstigerweise mehr Bildpunkte benutzt werden, aber das kostet Rechenzeit. Um die Rechenzeit der FF zu verkürzen, werden nicht alle Bildpunkte im Bild benutzt. Für das Leistungsspektrum werden nur 32x32 Bildpunkte verwendet, die innerhalb eines Fensters (62,62) - (372,372) im Abstand von 10 Bildpunkten rasterförmig verteilt sind. Für alle anderen Fokussierungsfunktionen wird das Fenster (0,0) - (511,484) benutzt. Dabei wird jede neunte Zeile verwendet. Insgesamt werden also $512 \times 54 = 27648$ Bildpunkte benutzt.

Während der Experimente sind alle Parameter der Kamera, z.B. die Blendenzahl, die Brennweite usw., festgelegt. Für alle Fokussierungsfunktionen werden folgende Experimente durchgeführt.

4.1 Der Einfluß der Beleuchtung

Ein Objekt wird sichtbar durch Reflexion, Brechung oder Streuung von Licht. Objekteigenschaften spielen dabei eine Rolle wie die Natur der Lichtquelle (Beleuchtungsverhältnisse). Während in "natürlichen" Szenen die Lichtverhältnisse meistens vorgegeben sind, hat man bei der Aufnahme wissenschaftlicher Objekte viele Möglichkeiten, die Beleuchtungsverhältnisse optimal zu gestalten. Die Aufnahmebedingungen stehen in enger Beziehung zu der geplanten Auswertung der Bilder.

Die Frage der Beleuchtungsverhältnisse besteht im einfachsten Fall aus zwei Teilaspekten (Bild 4.2). Das von einer oder mehreren Lichtquellen ausgestrahlte Licht trifft mit einer gewissen Beleuchtungsstärke auf dem Objekt auf. Dort wird es absorbiert, gebrochen und reflektiert.

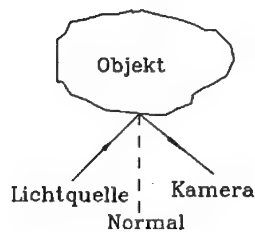


Bild 4.2: Beleuchtungsverhältnisse bei Objekten.

Der Bruchteil der in den Raumwinkel gestreuten Strahlen, die die Kamera treffen, bestimmt die Grauwerte, mit der das Objekt gesehen wird. Damit bestimmen sowohl die Beleuchtungsstärke als auch das Reflexionsverhalten die Grauwerte des Objektes.

In natürlichen Szenen hat man es in der Regel mit undurchsichtigen Oberflächen zu tun. In diesem Fall muß nur das Reflexionsverhalten von Oberflächen untersucht werden. Betrachtet man die von einer Lichtquelle ausgesandten Strahlen, so gibt es zwei Fälle der Betrachtungsweise. Erstens den Fall, daß in Richtung

der Lichtquelle beobachtet auf diese Strahlen Bezug genommen wird. Zweitens, daß diese Strahlen in bezug auf eine Fläche betrachtet werden, auf die das ausgestrahlte Licht auftrifft.

Unter der Bedingung der undurchsichtigen Oberfläche der Objekte werden folgende Untersuchungen durchgeführt:

a) Untersuchung des Einflusses auf die FF durch Veränderung der Beleuchtungsstärke.

Für eine gute FF ist der Einfluß der Beleuchtungsstärke unempfindlich; aber fast alle von der Kamera aufgenommenen Bilder sind durch das Rauschen verschmiert. Es ist klar, wenn die Beleuchtungsstärke groß ist, ist das Signal-zu-Rauschverhältnis (SNR) auch groß. Deshalb beeinflußt das Rauschen die FF nicht so stark.

In den Experimenten wird die FF bei verschiedenen Beleuchtungsstärken gemessen. Eine zusätzliche Lichtquelle wird benutzt, um die Beleuchtungsstärke zu verändern. Die Beleuchtungsstärke wird durch den Mittelwert des Bildes repräsentiert.

Die Fokussierungsfunktionen werden in drei Gruppen eingeteilt. Die erste Gruppe besteht aus den Fokussierungsfunktionen VAR, ENT, LS und TV. Die FF von dieser Gruppe ist durch die globale Bildverarbeitung zu bekommen. Die zweite Gruppe ist die SMD und die LO. Die Berechnung von SMD und LO braucht nur die arithmetischen Operationen mit den Nachbargrauwerten des jeweiligen Bildpunktes. Die dritte Gruppe mit der SL' ist die sogenannte punktbezogene FF.

Aufgrund der Ergebnisse der Experimente ist deutlich erkennbar, daß die Kurven der ersten Gruppe wenig durch die Beleuchtungsstärke beeinflußt werden. Bild 4.3a zeigt die Varianz der Grauwerte bei verschiedenen Beleuchtungsstärken, weil durch die globale Verarbeitung das Rauschen des Bildes unterdrückt wird. Aber die FF der zweiten und dritten Gruppe werden dagegen durch die Beleuchtungsstärke intensiv beeinflußt. Im Bild 4.3b hat die Kurve bei einer bestimmten Beleuchtungsstärke ihre unimodulare Form verloren. Die Gründe hierfür sind, daß durch die lokale Differentiation der Grauwerte nicht nur die Grauwertekante, sondern auch das Rauschen hervorgehoben werden.

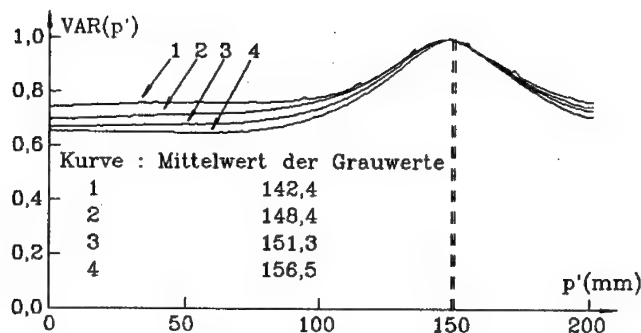


Bild 4.3a: Varianz bei verschiedenen Beleuchtungsstärken.

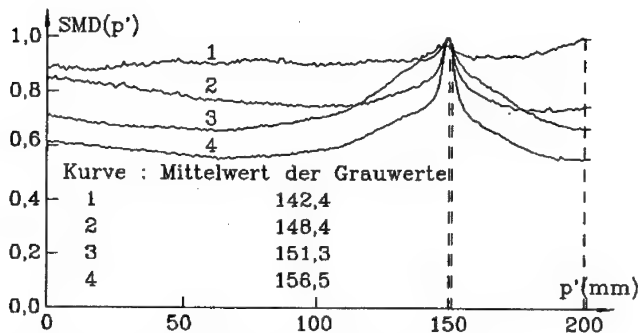


Bild 4.3b: SMD bei verschiedenen Beleuchtungsstärken.

- b) Untersuchung des Einflusses auf die FF durch die Veränderung der Beleuchtungsrichtung.

Der Einfluß der Beleuchtungsrichtung wird durch das Reflexionsverhalten der Oberfläche des Objektes bestimmt. Wenn die Szene eine ebene Oberfläche besitzt, stehen die Normalen aller Flächenelemente unter einer bestimmten Richtung zu dem einfallenden Strahl. Ein Flächenelement erhält nur Licht von der Lichtquelle. In dieser Situation spielt die Beleuchtungsrichtung nur eine geringe Rolle (Bild 4.4). Aber wenn die Szene aus dreidimensionalen Gebilden besteht, stehen die Normalen der Flächenelemente in verschiedenen Richtungen zu dem einfallenden Strahl. So erhält ein Flächenelement nicht nur Licht von den Lichtquellen direkt, sondern auch indirekt von allen anderen Flächenelementen. Deshalb beeinflussen die

Lage und das Reflexionsverhalten aller Oberflächenelemente in einer Szene den Grauwert eines Oberflächenelementes. In diesem Fall spielt die Beleuchtungsrichtung eine wichtige Rolle für die FF. Wie Bild 4.5 zeigt, ist die FF eine regellose Kurve, wenn die Lichtquelle direkt von vorn auf die Szene gerichtet ist.

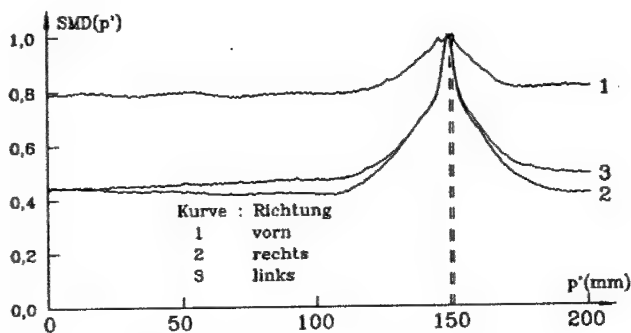


Bild 4.4: SMD in verschiedenen Beleuchtungsrichtungen.

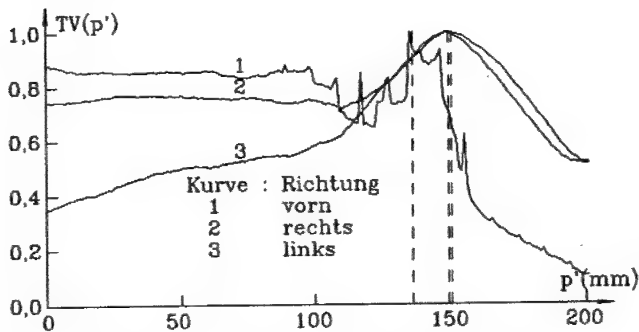


Bild 4.5: TV in verschiedenen Beleuchtungsrichtungen.

Auch wenn die Positionen der Objekte und Lichtquellen und die Reflexionseigenschaften bekannt sind, ist es trotzdem sehr schwierig, die Grauwerte durch Lösung eines gekoppelten Gleichungssystems, in dem die entsprechende Anzahl Oberflächenelemente enthalten ist, zu ermitteln. Deshalb ist es noch unmöglich, den Einfluß der Beleuchtungsrichtung auf der FF quantitativ zu analysieren.

c) Untersuchung des Einflusses auf die FF durch Veränderung der Beleuchtungsfarben.

Die Lichtquelle strahlt weißes Licht aus. Weißes Licht ist im streng physikalischen Sinne ein aus allen Wellenlängen des sichtbaren Spektralbereiches energieggleich gemischtes Licht. Die Wellenlängen liegen für sichtbares Licht zwischen $\lambda = 380 \text{ nm}$ und $\lambda = 780 \text{ nm}$. Die Energieverteilung auf die einzelnen Frequenzintervalle bestimmt die Farbzusammensetzung und den Farbeindruck. Das menschliche Auge spricht auf Licht unterschiedlicher Wellenlänge verschieden stark an.

Filter ändern die spektrale Energieverteilung einer durchgehenden Strahlung. Als einfachstes Filter sind die Farbgläser anzuführen, die aufgrund ihrer selektiven Absorption einen Teil der Strahlung wegfiltern und die Lichtstärke läßt durch die Farbgläser nach. Für die Photographie dienen Filter in erster Linie zur richtigen Wiedergabe der Farbhelligkeiten in den entsprechenden Grauwerten. Bei starker Beleuchtung werden Filter verwendet, um die kontrastschwächende Strahlung zu absorbieren. Damit wird die Bildqualität verbessert.

In den Experimenten werden drei Farbgläser mit den Farben Indigo, Grün und Rot benutzt. Die Wellenlängen von Indigo, Grün und Rot sind $430,8 \text{ nm}$, $527,0 \text{ nm}$ und $686,3 \text{ nm}$. Werden die mit Indigo gefärbten Gläser benutzt, ist der enge Spektralbereich in der Nähe von $430,8 \text{ nm}$ durchlässig, die anderen Spektralbereiche sind dagegen undurchlässig. Bei den anderen Farben entsprechend.

Die Ergebnisse zeigen, daß - wenn die Beleuchtung nicht zu stark ist - der Einfluß der Farbgläser auf die FF nur schwer erkennbar ist (Bild 4.6).

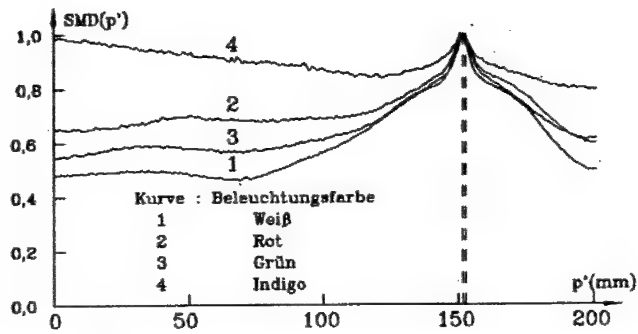


Bild 4.6: SMD bei verschiedenen Beleuchtungsfarben.

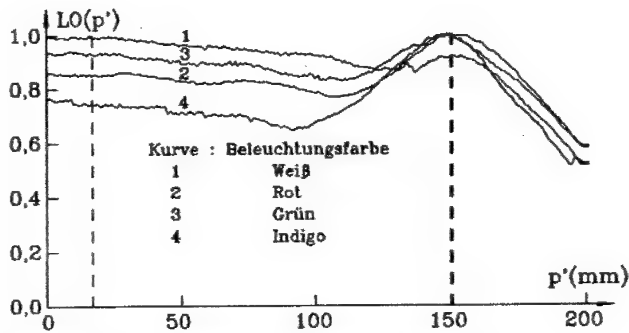


Bild 4.7: LO bei verschiedenen Beleuchtungsfarben.

Ist die Beleuchtung stark, wird der Kontrast im Bild abgeschwächt. Durch die Benutzung der Farbgläser kann die Bildqualität erhöht werden, damit wird die Kurve der FF deutlich verbessert (Bild 4.7).

4.2 Der Einfluß von Mustern des Objektes

Um den Einfluß auf die FF durch Muster zu prüfen, werden folgende Muster vorbereitet (Bild 4.8).

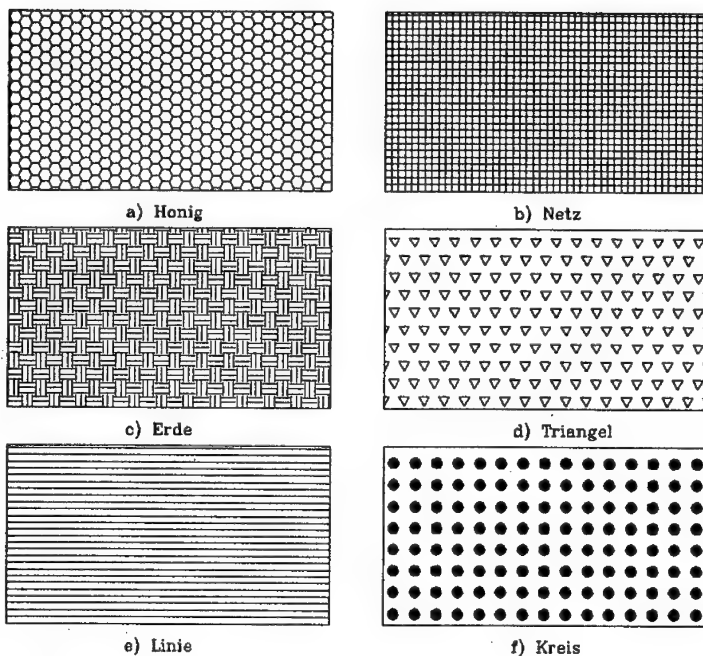


Bild 4.8: Verschiedene Muster.

Alle sechs Muster sind künstlich, sie beinhalten unterschiedliche Texturen. Die Ergebnisse der Experimente zeigen, daß alle Fokussierungsfunktionen außer SMD und ENT ein gutes Anpassungsvermögen an Muster haben.

Die Fokussierungsfunktion ENT ist kein zuverlässiges Fokussierungskriterium. Es liefert zwar in den meisten Fällen Ergebnisse, die mit den Ergebnissen der anderen FF im Einklang stehen, versagt jedoch vollständig bei Bildern, in denen wenige Grauwerte gehäuft auftreten.

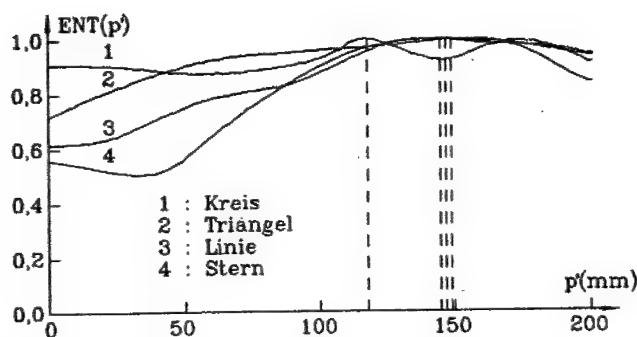


Bild 4.9: ENT von verschiedenen Mustern.

Die Kurve 1 in Bild 4.9 ist die ENT des Musters Kreis. In der besten Fokussierungsposition liefert ENT nicht das Maximum, sondern das Minimum. Außerdem ist die Kurve von ENT zu flach. In der Nähe der besten Fokussierungsposition verändert sich die ENT wenig. Deshalb ist die Genauigkeit von ENT niedrig. Weil die ENT in bestimmten Fällen vollständig versagt und gegenüber den anderen FF keine Vorteile aufweist, soll sie bei den weiteren Messungen nicht mehr verwendet werden.

Die SMD1 als FF liefert keine Informationen über den Fokussierungszustand, wenn die Szene aus einem Linienmuster besteht. SMD1 extrahiert nur den Kantengradient entlang der x-Richtung. Wenn die SMD3 als FF benutzt wird, kann dieses Problem allerdings vermieden werden (Bild 4.10).

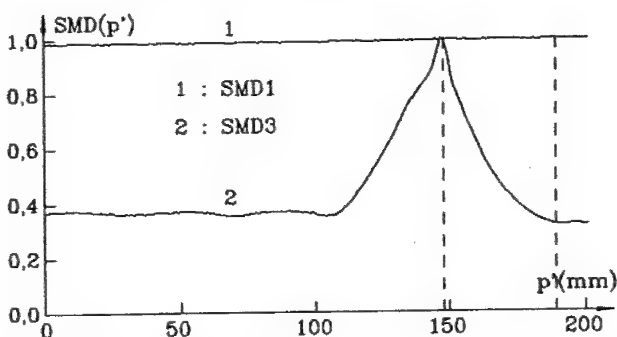


Bild 4.10: SMD eines Linienmusters.

4.3 Die Reaktion der Fokussierungsfunktionen bei texturfreien Bildern

Alle Fokussierungsfunktionen basieren auf der Fluktuation der Grauwerte des Bildes. Das Bild von einem texturfreien Objekt, z.B. ein weißes Papier, ist homogen, d.h. in diesem Bild tritt nur ein bestimmter Grauwert auf. In dieser Situation sind alle Leistungsspektren, außer dem Leistungsspektrum vom Ursprung, also $u = 0$ und $v = 0$, Null. Deshalb sind die aufgeführten Fokussierungsfunktionen ungeeignet zur Fokussierung eines texturfreien Objektes. Wenn das optische System ideal ist, d.h. das Bild nicht durch Rauschen verschmiert ist, ist die Fokussierungsfunktion eines texturfreien Bildes konstant. Durch die Fokussierungsfunktion eines texturfreien Bildes kann man deshalb den Einfluß von Rauschen direkt bemerken und auch extrahieren..

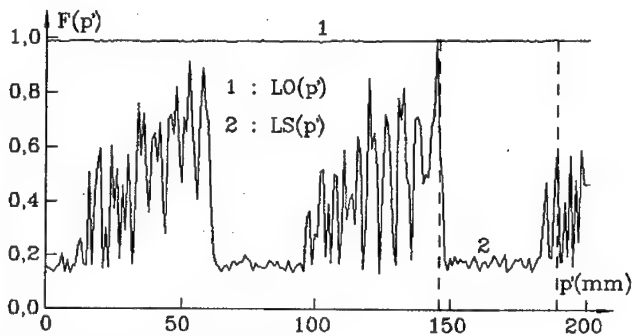


Bild 4.11: FF des texturfreien Bildes.

Bild 4.11 zeigt $LO(p')$ und $LS(p')$ des texturfreien Bildes. Die anderen Fokussierungsfunktionen sind ähnlich wie $LO(p')$ und nicht so stark durch Rauschen beeinflusst. Weil das Leistungsspektrum als FF wenige Bildpunkte benutzt, ist der Einfluß von Rauschen deutlich bemerkbar. Den Einfluß des Rauschens kann man auch durch die Fokussierungsfunktionen $LO(p')$ und $LS(p')$ beim Muster Honig bemerken (Bild 4.12).

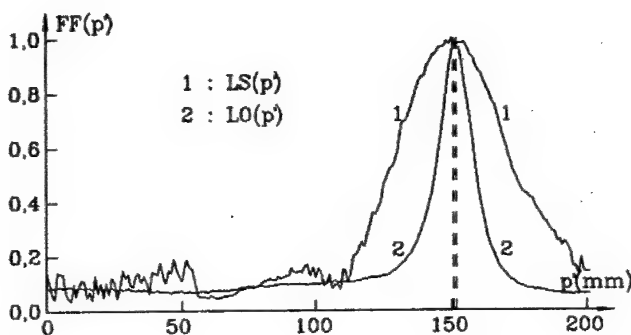


Bild 4.12: FF vom Muster Honig.

Im wirksamen Bereich der Kamera ist die Kurve von $LS(p')$ gut, aber im unwirksamen Bereich hat $LS(p')$ große Schwingungen.

4.4 Die Fokussierungsfunktionen bei farbigen Mustern

Das Ziel dieses Experiments ist die Untersuchung des Reflexionsgrads von farbigen Mustern. Das Objekt ist das Muster Quadrat auf dem weißen Papier (Hintergrund). Der Hintergrund bleibt unverändert, die Farben des Musters sind unterschiedlich:

- a) schwarz,
- b) blau,
- c) gelb,
- d) rot.

In dem Experiment wird die Beleuchtung nicht verändert. Es ist bekannt, daß der Grauwert $g(x,y)$ sich aus dem Produkt der Beleuchtungsstärke $i(x,y)$ und des Reflexionsgrads vom Objekt $l(x,y)$ ergibt:

$$g(x,y) = i(x,y) \cdot l(x,y)$$

Der Reflexionsgrad des Hintergrundes ist sehr groß, deshalb ist der Hintergrund im Bild immer hell. Wenn der Reflexionsgrad des Musters klein ist, also das Bild des Musters dunkel bleibt, ist der Kontrast des Bildes stark. Deshalb erhält man eine gute Kurve der FF.

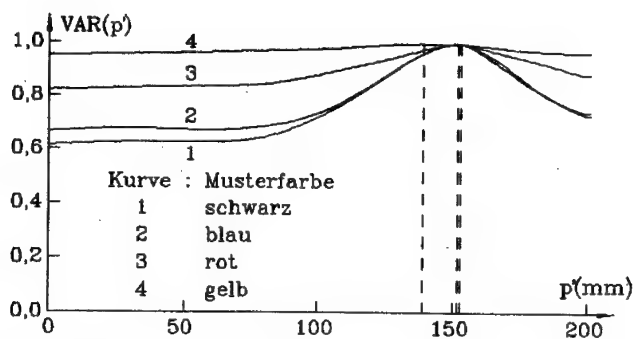


Bild 4.13: VAR von farbigen Mustern.

In Bild 4.13 ist Kurve 1 die VAR vom schwarzen Muster. Sie ist die günstigste Kurve zur Fokussierung. Die schwarze Farbe kann wenig Licht reflektieren. Deswegen ist der Kontrast des Bildes stark. Wenn das Muster blau ist, kann man auch eine gute Kurve der VAR bekommen. Kurve 1 und Kurve 2 überlappen sich im wirksamen Bereich der Kamera fast. Kurve 3 ist ziemlich flach, mit Hilfe dieser Kurve kann die beste Fokussierungsposition schwer gefunden werden. Kurve 3 ist die VAR des roten Musters. Durch dieses Experiment bemerkt man, daß der Reflexionsgrad von Rot größer als der von Schwarz und Blau ist. Ist das Muster gelb, verliert das Bild seinen Kontrast total. In Kurve 4 ist VAR im ganzen Fokussierungsbereich fast unverändert. Kurve 4 enthält keine Information über den Fokussierungszustand der Kamera.

Die Ergebnisse des Experiments zeigen, wenn eine farbige Szene durch Digitalisierung in ein monochromatisches Bild umgewandelt wird, gehen viele Informationen der Szene verloren.

4.5 Die Fokussierungsfunktionen bei unscharfen Konturen

Manchmal hat die zu fokussierende Szene eine unscharfe Kontur, z.B. besteht die Szene aus Objekten mit unscharfen Bildern. Wie gezeigt wurde, besitzen die unklaren Bilder geringere Leistungsspektren im hohen Frequenzbereich. Die Fokussierung solcher Szenen ist natürlich schwer. Die Buchstaben in Bild 4.14a sind fokal. Bild 3.14b zeigt die gleichen Buchstaben mit unscharfen Konturen, man kann aber noch identifizieren. Bild 4.14c und d sind total unscharf, man kann die Buchstaben nicht erkennen.

svensk politik, Stockholm 1984, d
verfahren. Andere Autoren deuten si
te und Vorurteile hin; vgl. Rune Pre
i offentliga beslutsprocesser, in: 1
Utvärdering av offentlig verksamhet
ingens roll i samhället, Stockholm 11
s: kunskap, praktik och etik i offent
1989. Als Übersicht vgl. Evert Vedun
och förvaltning, Lund 1991.

a)

svensk politik, Stockholm 1984, d
verfahren. Andere Autoren deuten si
te und Vorurteile hin; vgl. Rune Pre
i offentliga beslutsprocesser, in: 1
Utvärdering av offentlig verksamhet
ingens roll i samhället, Stockholm 11
s: kunskap, praktik och etik i offent
1989. Als Übersicht vgl. Evert Vedun
och förvaltning, Lund 1991.

b)

svensk politik, Stockholm 1984, d
verfahren. Andere Autoren deuten si
te und Vorurteile hin; vgl. Rune Pre
i offentliga beslutsprocesser, in: 1
Utvärdering av offentlig verksamhet
ingens roll i samhället, Stockholm 11
s: kunskap, praktik och etik i offent
1989. Als Übersicht vgl. Evert Vedun
och förvaltning, Lund 1991.

c)

svensk politik, Stockholm 1984, d
verfahren. Andere Autoren deuten si
te und Vorurteile hin; vgl. Rune Pre
i offentliga beslutsprocesser, in: 1
Utvärdering av offentlig verksamhet
ingens roll i samhället, Stockholm 11
s: kunskap, praktik och etik i offent
1989. Als Übersicht vgl. Evert Vedun
och förvaltning, Lund 1991.

d)

Bild 4.14: Bilder mit unscharfen Konturen.

Die Meßergebnisse demonstrieren, wenn die Buchstaben erkennbar sind, zeigen die Kurve 1 und die Kurve 2 eine gute Übereinstimmung und erreichen bei gleicher Fokussierungsposition das Maximum. Aber Kurve 2 ist deutlich flacher als

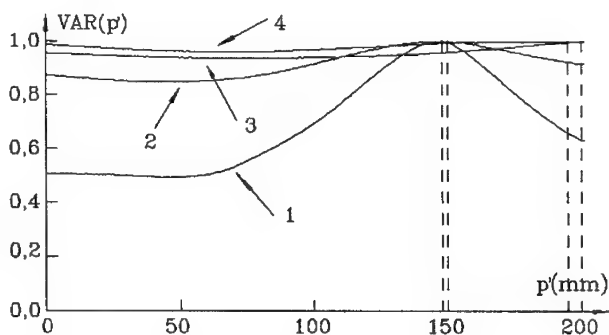


Bild 4.15: VAR mit unscharfen Konturen.

Kurve 1 (Bild 4.15). Weil Bild 4.14c und Bild 4.14d zu unscharf sind, sind in Kurve 3 und in Kurve 4 die Fähigkeit zur Fokussierung verlorengegangen. Es ist unmöglich, diese beiden Bilder mit Hilfe der FF zu fokussieren.

4.6 Der Einfluß durch starke Reflexion

Bei spiegelnden Oberflächen wird einfallendes Licht nur in einer Richtung reflektiert. Viele Objekte enthalten einen Teil dieser Reflexionsart, die in Form von Glanzlichtern in den Bildern sichtbar ist. Diese Bildteile sind für die Bildverarbeitung besonders problematisch, da sie selbst bei geringer Drehung des Objektes sich auf diesem verschieben und daher ihr Aussehen verändern. Die Spiegelung erfolgt, wenn die Rauigkeit klein gegenüber den Wellenlängen des einfallenden Lichtes ist. Ideale glatte Oberflächen sind nur durch Kristallspaltflächen realisiert.

Als Spiegelflächen benutzt man Metallflächen, die meist aus Aluminium oder Silber bestehen. Im vorliegenden Experiment wird Glanzpapier als Spiegelfläche benutzt. Der Reflexionsgrad des Glanzpapiers ist fast 1, d.h. das Glanzpapier kann das einfallende Licht zu fast 100 % reflektieren. In dieser Situation ist das reflektierte Licht zu stark. Wenn das Glanzpapier keine Falte besitzt, ist das Bild vom Glanzpapier homogen. Deshalb kann die Fokussierungsfunktion ein solch homogenes Bild nicht fokussieren. Wenn auf der Oberfläche des Glanzpapiers Falten auftreten, d.h. die Oberfläche des Glanzpapiers uneben ist, wird das Licht in verschiedenen Richtungen reflektiert. Das Bild des Glanzpapiers ist nicht mehr homogen.

Wenn das Glanzpapier mit Falten mit den globalen Fokussierungsfunktionen, z. B. VAR, zu fokussieren ist, wird die Kamera die falsche Fokussierungsposition finden. Dieses Phänomen kann man durch die Ergebnisse des Experiments aufdecken. Kurve 1 in Bild 4.16 ist VAR von Glanzpapier. Das Maximum ist nicht in der besten Fokussierungsposition $p' = 150 \text{ mm}$, sondern in der Position $p' = 120 \text{ mm}$ aufgetreten. In der besten Fokussierungsposition tritt die VAR als ein lokales Minimum auf.

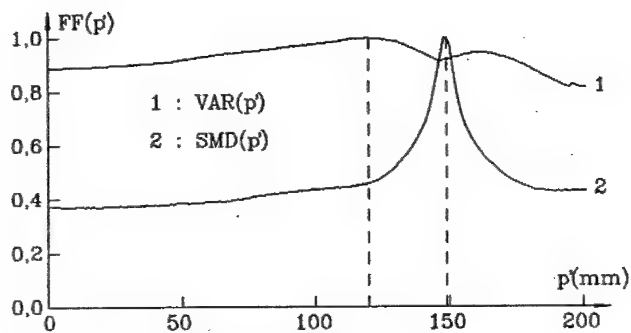


Bild 4.16: Fokussierungsfunktionen vom Glanzpapier.

Kurve 2 ist die lokale Fokussierungsfunktion SMD. Das Maximum in der Kurve 2 ist genau in der besten Fokussierungsposition aufgetreten. Um dieses Phänomen zu erklären, werden die Bilder des Glanzpapiers in der Fokussierungsposition $p' = 120$ mm und $p' = 150$ mm in Bild 4.17 gezeigt.



a) $p' = 120$ mm



b) $p' = 150$ mm

Bild 4.17: Bild des Glanzpapiers mit Falten.

Das Bild von $p' = 120$ mm besteht aus zwei homogenen Bereichen. Ein Bereich ist schwarz, der andere im Gegensatz dazu ist weiß. In diesem Bild ist der Kontrast am stärksten. Deshalb erreicht die VAR in dieser Position ihr Maximum, aber für die lokale Fokussierungsfunktion SMD ist ihr Wert innerhalb des homogenen Bereiches Null. Nur in dem Übergang zwischen beiden homogenen Bereichen ist der Wert der SMD von Null verschieden. Deswegen hat SMD in der Position $p' = 120$ mm ihren kleinsten Wert. Bild 4.17b zeigt das Bild vom Glanzpapier in der besten Fokussierungsposition. In diesem Bild kann man die Textur des Glanzpapiers erkennen.

4.7 Fokussierungsfunktionen bei dreidimensionalen Gebilden

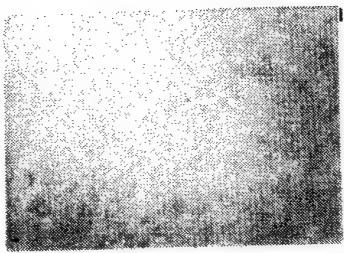
In den bisherigen Experimenten sind die Oberflächen der Objekte immer eben. Die ebenen Oberflächen sind von einer ideal glatten Fläche nur geringfügig abgewichen und haben eine niedrige Rauigkeit. Wegen der ebenen Oberfläche beleuchtet das Licht der Lichtquelle die Szene überall. Deshalb treten im Bild keine Schatten auf. Wenn die Beleuchtungsstärke als konstant angenommen wird, d.h. eine Punktlichtquelle im Unendlichen existiert, ist der Grauwert in der Position (x,y) nur vom Reflexionsgrad in diesem Punkt abhängig.

In diesem Experiment werden die Fokussierungsfunktionen bei dreidimensionalen Gebilden gemessen. Als zu untersuchende Objekte werden folgende dreidimensionalen Gebilde gewählt:

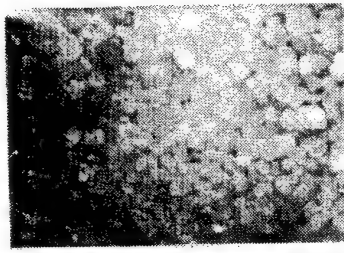
- Mehl,
- Sand,
- Sesam,
- Haselnuß.

Bild 4.18a ist das Bild von Mehl, dessen räumliche Dimension im Vergleich zu anderen Gebilden sehr klein ist. Im Bild ist keine Textur zu bemerken. Das Bild besitzt einen sehr schwachen Kontrast. Die Oberfläche der Szene mit Mehl ist ziemlich flach, deswegen treten in dem Bild auch keine Schatten auf.

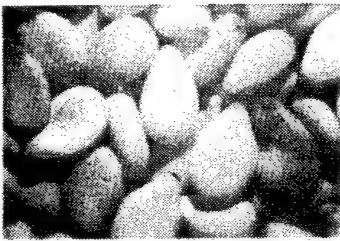
Das Bild von Sand in Bild 4.18b ist ein texturreiches Bild. Da die räumliche Dimension des Sandes noch nicht groß genug ist, hat auch dies Bild keinen Schatten. Der Kontrast von Bild 4.18b ist zwar nicht so stark, die Grauwerte der benachbarten Bildpunkte verändern sich aber deutlich. Deshalb sind die lokalen Fokussierungsfunktionen günstig zur Fokussierung der Szene mit Sand.



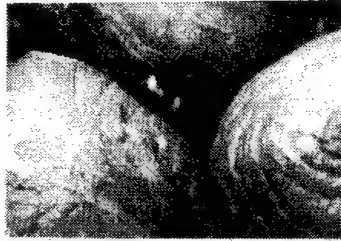
a)



b)



c)



d)

Bild 4.18: Bilder verschiedener dreidimensionaler Gebilde.

Die räumliche Dimension von Sesam ist im Vergleich zum Sand sehr groß. In Bild 4.18c kann man den Schatten deutlich sehen. Der Schatten ist die Abbildung des Lichtes nicht oder weniger beleuchteter Teile einer Fläche, hervorgerufen durch undurchsichtige Körper, die sich im Strahlengang befinden. Das Schattengebiet wird durch die Gestalt des Körpers und die Gestalt der Lichtquelle bedingt. Deshalb spielt die Beleuchtungsrichtung eine große Rolle für das Bild von dreidimensionalen Gebilden. Wenn das Bild einen Schatten hat, tritt eine Kante zwischen den Schattengebieten und den beleuchteten Gebieten auf. Diese Kante kann zwar die Grauwertfluktuation im Bild verstärken, besitzt aber keine Information über den Fokussierungszustand. Wenn das Schattengebiet im Bild sehr groß ist, kann die beste Fokussierungsposition mit Hilfe der FF nicht gefunden werden. Bevor ein dreidimensionales Gebilde zu fokussieren ist, muß deshalb die Beleuchtungsrichtung so ausgerichtet werden, daß das Schattengebiet im Bild so klein wie möglich wird.

Die Oberfläche von Sesam ist ziemlich glatt. Die Oberfläche im Bild ist ein homogener Bereich, deshalb tritt die Kante im Bild nur zwischen zwei Körnern auf.

Das letzte Gebilde ist die Haselnuß. Die räumliche Dimension der Haselnuß ist im Vergleich zu den anderen Gebilden sehr groß. Deshalb hat Bild 4.18d große Schattengebiete. Die Oberfläche der Haselnuß ist sehr rauh. Die große Rauigkeit der Oberfläche kann zur Fokussierung des Objektes hilfreich sein.

Bild 4.19a zeigt die globale Fokussierungsfunktion VAR. Wenn die Szene aus Mehl besteht, ist die Grauwertfluktuation sehr klein. Kurve 1 ist die VAR vom Mehl und hat keine unimodulare Eigenschaft mehr, d.h. die globalen Fokussierungsfunktionen sind ungeeignet zur Fokussierung dieses kontrastarmen Bildes.

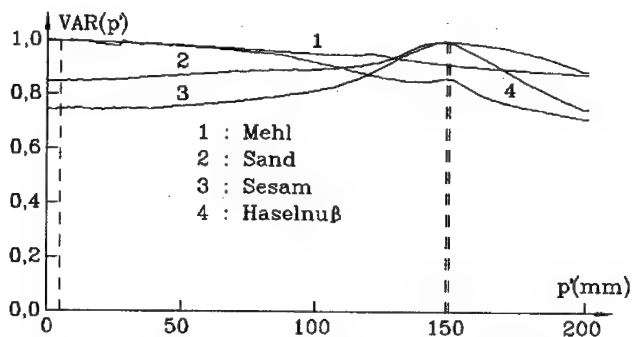


Bild 4.19a: VAR bei verschiedenen dreidimensionalen Gebilden.

Kurve 4 ist die VAR der Haselnuß. Zwar verläuft die Kurve in die beste Fokussierungsposition eines lokalen Maximums, aber das globale Maximum befindet sich am Anfang des Fokussierungsbereiches. In Kurve 2 und Kurve 3 erscheint das Maximum in der besten Fokussierungsposition, aber diese beiden Kurven sind zu flach.

Die lokalen Fokussierungsfunktionen können die Bilder mit schwachem Kontrast problemlos fokussieren. Die SMD für die vier dreidimensionalen Gebilden ist in Bild 4.19b dargestellt.

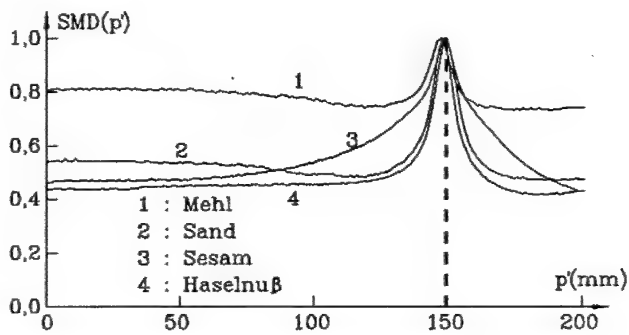
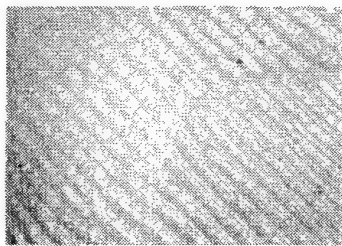
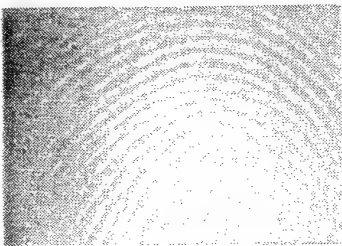


Bild 4.19b: SMD bei verschiedenen dreidimensionalen Gebilden.

Die Maxima von allen vier Kurven in Bild 4.19b befinden sich genau in der besten Fokussierungsposition.

Die Fingerabdruckerkennung spielt in der Kriminologie eine wichtige Rolle. Der Fingerabdruck ist ein typisch dreidimensionales Gebilde. In Bild 4.20a kann man die Textur davon deutlich erkennen.



a) Fingerabdruck

b) Haut der Handfläche

Bild 4.20: Bilder der menschlichen Haut.

Die Grauwerte in diesem Bild verändern sich regelmäßig. Bild 4.20b zeigt das Bild der Haut einer Handfläche. Man kann die Richtung der Grauwerte deutlich erkennen. Beide Bilder sind arm an Kontrast, deshalb sind nur die lokalen Fokussierungsfunktionen zu benutzen, um die Haut zu fokussieren. Die Ergebnisse sind in Bild 4.21 dargestellt.

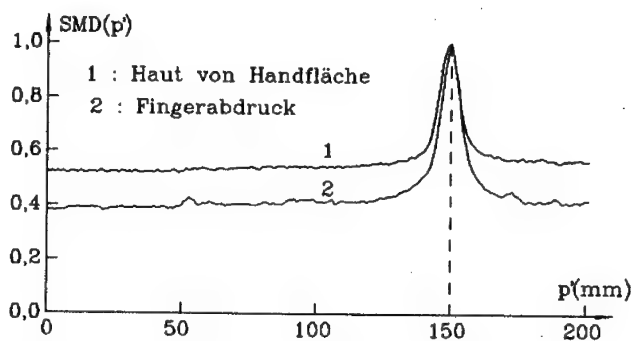
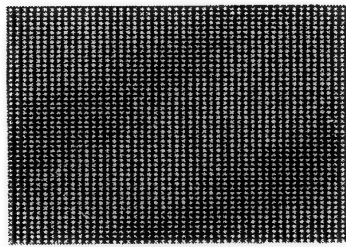


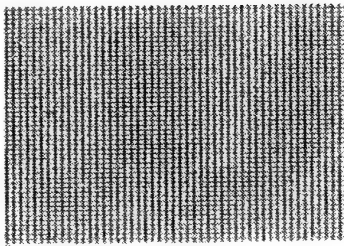
Bild 4.21: SMD der menschlichen Haut.

4.8 Fokussierungsfunktionen von schwarz-weiß Mustern unterschiedlicher Grautönung

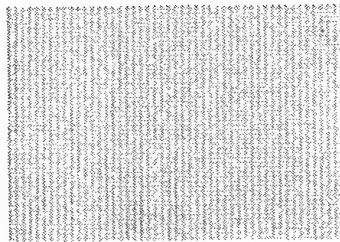
Dieses Experiment ist zur Untersuchung des Einflusses des Kontrastes vom Objekt entworfen worden. Die Objekte sind die Gitter auf dem weißen Papier. Es ist klar, daß der Grauwert des weißen Papiers sehr groß ist. Der Grauwert der Gitter wird verändert, damit man die schwarz-weiß Muster mit unterschiedlicher Grautönung bekommen kann. Die Grauwerte der Gitter haben drei Stufen, nämlich 0, 100 und 200.



a)



b)



c)

Bild 4.22: Schwarz-weiß Muster unterschiedlicher Grautönung.

Wenn die Grauwerte der Gitter Null sind, hat das Bild den stärksten Kontrast. Im Gegensatz dazu ist beim Grauwert 200 der Kontrast des Musters am schwächsten.

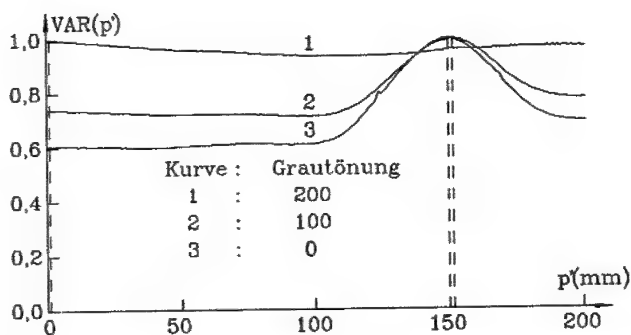


Bild 4.23: VAR von schwarz-weiß Mustern unterschiedlicher Grautönung.

Die Ergebnisse sind klar. Je stärker der Kontrast des Objektes ist, desto besser für die Fokussierung. Wenn der Kontrast des Objektes zu schwach ist, kann das Objekt nicht mit Hilfe der FF fokussiert werden.

4.9 Zusammenfassung

Es wurden Experimente zur Untersuchung der Eigenschaften von verschiedenen Fokussierungsfunktionen durchgeführt. Durch diese Experimente werden die Vorteile und Nachteile der Fokussierungsfunktionen unter folgenden Aspekten beurteilt:

1. Kompliziertheit der Fokussierungsfunktionen.

Alle Fokussierungsfunktionen außer der Fokussierungsfunktion LS haben einfache Algorithmen. Zur Berechnung des Leistungsspektrums vom Bild ist eine zweidimensionale Fouriertransformation durchzuführen. Um Rechenzeit zu sparen, wird ein Algorithmus der schnellen Fouriertransformation (FFT) verwendet. Die FFT ist ein relativ aufwendiger Algorithmus.

Durch die Ableitung der Varianz als Fokussierungsfunktion wird die Störung des optischen Systems stark hervorgehoben. Um dieses Problem zu lösen, wird ein digitales Filter benutzt. Die Benutzung des digitalen Filters verursacht allerdings eine Verzögerung der besten Fokussierungsposition. Diese Verzögerung kann zwar vorhergesagt und durch das Programm kompensiert werden, aber das Programm zur Kompensierung der Verzögerung und der Algorithmus des digitalen Filters verkomplizieren das Fokussierungsverfahren. Deshalb ist die Ableitung der Varianz als FF gegenüber der Varianz selbst ein komplizierterer Algorithmus.

2. Die Rechenzeit der Fokussierungsfunktionen.

Die Rechenzeit ist nicht so wichtig wie die Genauigkeit der FF. Aber für eine gute FF ist die Rechenzeit sehr kurz. Die Rechenzeit der verschiedenen FF sind in folgender Tabelle gezeigt.

FF	VAR	ENT	SMD	LS	SL'	TV	LO
Rechen-Zeit(s)	0,24	0,23	0,32	1,00	0,68	0,20	0,36

Tab. 4.1: Rechenzeit der verschiedenen Fokussierungsfunktionen.

Die Rechnung für TV braucht nur einfache arithmetische Entscheidungen durchzuführen. Deshalb ist die Rechenzeit von TV am kürzesten.

Um die VAR und die ENT zu berechnen, braucht der Computer 0,24 s bzw. 0,23 s. Da die Rechnung von VAR und ENT das Histogramm des Bildes benutzt hat, ist die Rechenzeit sehr kurz.

LS ist eine zeitraubende Fokussierungsfunktion. Obwohl die Anzahl der benutzten Bildpunkte geringer als bei anderen Fokussierungsfunktionen ist, beträgt die Rechenzeit für LS 1 s.

Weil für die Berechnung von SL' die Gleitkommazahlen mit doppelter Genauigkeit benötigt werden, ist die Rechenzeit ziemlich lang (0,68 s).

3. Anpassungsvermögen der FF an Objekte.

In den Experimenten wurden sowohl Objekte von künstlichen Mustern mit regelmäßigen Texturen als auch Objekte von natürlichen Mustern (dreidimensionale Gebilde) mit überwiegend unregelmäßigen Texturen verwendet. Die Ergebnisse zeigen, daß die Fokussierungsfunktion SMD3 und LO die besten Anpassungsvermögen an die verschiedenen Objekte haben. SMD1 ist ungeeignet zur Fokussierung des Objektes mit der Kante entlang der y-Richtung. SMD2 kann das Objekt mit der Kante entlang der x-Richtung nicht fokussieren.

Die Fokussierungsfunktion ENT versagt vollständig bei Bildern, in denen wenige Grauwerte gehäuft auftreten. Die globalen Fokussierungsfunktionen VAR, LS, TV sind schlecht zur Fokussierung der Objekte mit schwachem Kontrast. Diese drei Fokussierungsfunktionen sind ungünstig zur Fokussierung der dreidimensionalen Gebilde.

4. Genauigkeit der Fokussierungsfunktionen.

Um die Genauigkeit der FF zu beurteilen, wurde folgendes Experiment durchgeführt: Die Kamera fokussiert ein bestimmtes Objekt mit allen Fokussierungsfunktionen. Für jede FF wird der Fokussierungsprozeß 15mal durchgeführt. Inzwischen verändert sich die Beleuchtungsstärke. Jedesmal wird die gefundene beste Fokussierungsposition p'_{opt} notiert. Anschließend werden die durchschnittliche beste Fokussierungsposition und die mittlere Abweichung der besten Fokussierungsposition berechnet. Die Ergebnisse sind in Tabelle 4.2 dargestellt.

FF	p'_d (mm)	p'_{ma} (mm)	p'_b (mm)
VAR	151,14	1,187	150-153
ENT	153,86	8,450	139-162
SMD	151,30	0,499	151-152
LS	151,86	2,776	147-156
SL	151,75	2,026	148-156
TV	151,12	1,414	149-154
LO	152,23	0,973	150-153
DVAR	151,92	1,272	149-153

Tab. 4.2: Genauigkeit der Fokussierungsfunktionen.

In Tabelle 4.2 ist p'_d die durchschnittliche beste Fokussierungsposition; sie läßt sich mit folgender Gleichung berechnen

$$p'_d = \frac{1}{15} \cdot \sum_{i=1}^{15} p'^i_{opt}$$

p'_{ma} ist die mittlere Abweichung von der besten Fokussierungsfunktion

$$p'_{ma} = \sqrt{\frac{1}{15} \cdot \sum_{i=1}^{15} [p'^i_{opt} - p'_d]^2}$$

p'_b ist der Bereich der besten Fokussierungsfunktionen.

Für eine gute Fokussierungsfunktion ist die beste Fokussierungsposition immer in der gleichen Position. Aber der Einfluß des Objektes und des Rauschens auf die Fokussierungsfunktion sind immer vorhanden, wenn das optische System nicht aberrationsfrei ist, was bei jedem realen optischen System der Fall ist. Deshalb befindet sich die beste Fokussierungsfunktion nicht in einer Position. Solange sich die Fokussierungsposition innerhalb der Schärfentiefe befindet, wird diese Position als richtige Fokussierungsposition behandelt. Wegen der Auflösung des menschlichen Auges ist die Fokussierungsposition zwischen 149 mm und 152 mm die richtige Fokussierungsposition im beschriebenen System.

Für alle Fokussierungsfunktionen außer der ENT befinden sich die durchschnittlich besten Fokussierungspositionen p'_d innerhalb dieses Intervalls (149 mm - 152 mm). Weil die Kurve von ENT zu flach ist, verteilt sich die beste Fokussierungsfunktion auf einen breiten Bereich. Deshalb kann ENT das Objekt nur grob fokussieren: Normalerweise, wenn $p'_{ma} > 2$ mm ist, hat die Fokussierungsfunktion eine geringere Genauigkeit. Die p'_{ma} von ENT, LS und SL sind größer als 2 mm. Die Kurve von LO hat eine schärfere Spitze, deshalb ist die Genauigkeit von LO am besten. Die Fokussierungsfunktionen VAR, SMD und TV haben auch sehr gute Genauigkeiten.

5. Zuverlässigkeit der Fokussierungsfunktionen.

Die gute Fokussierungsfunktion hat eine starke Störfestigkeit. Der Einfluß äußerer Bedingungen für die Fokussierung ist nicht empfindlich, aber nur VAR und SMD gehören zur Klasse solcher Fokussierungsfunktionen. Für ENT spielen die Beleuchtungsstärke und die Beleuchtungsrichtung eine große Rolle; LS, TV, SL und LO sind auch sehr empfindlich gegenüber der Beleuchtungsstärke. Wenn die Beleuchtungsstärke zu klein ist, werden die Objekte falsch fokussiert.

6. Kurven der Fokussierungsfunktionen.

Eine gute Fokussierungsfunktion hat folgende Eigenschaften:

- a. Die Kurve der FF ist glatt. Das bedeutet, daß der Einfluß des Rauschens klein ist.

- b. In der Nähe der besten Fokussierungsposition verändert sich die FF sehr schnell. Deshalb ist die Genauigkeit der FF hoch.
- c. Die Breite des wirksamen Bereiches der Kamera ist groß, damit die Kamera in einem großem Arbeitsbereich fokussieren kann.

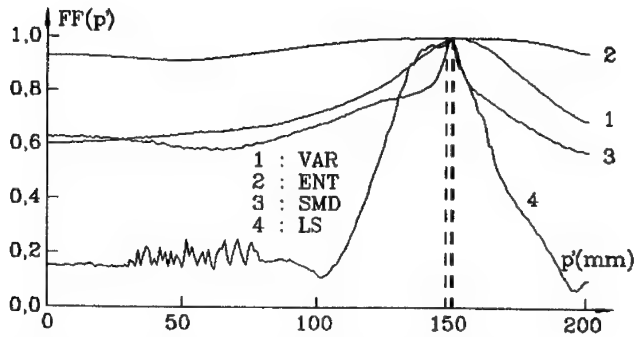


Bild 4.24: Kurven verschiedener Fokussierungsfunktionen.

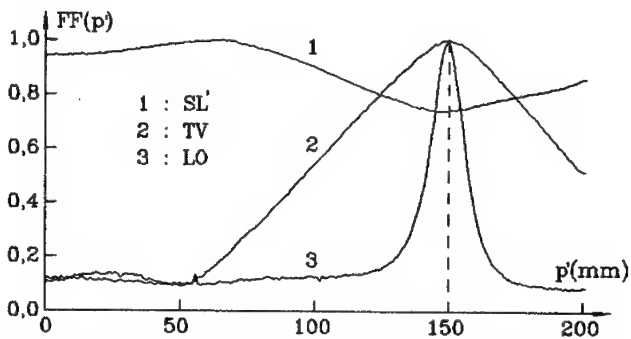


Bild 4.25: Kurven verschiedener Fokussierungsfunktionen.

Bild 4.24 und Bild 4.25 zeigen die Kurven von allen FF. Als Objekt dient das Muster Honig. Die Kurven von VAR, SMD und TV haben alle oben erwähnten Eigenschaften. Die Kurve von ENT ist zwar glatt, aber die Spitze ist zu flach.

Weil die Berechnung des Leistungsspektrums wenige Bildpunkte benutzt, ist der Einfluß des Rauschens sehr groß, besonders im unwirksamen Bereich der Kamera.

Die Fokussierungsfunktion SL' hat in der besten Fokussierungsposition kein Maximum, sondern ein Minimum. Die Kurve von SL' ist auch flach.

Die Kurve von LO verändert sich in der Nähe der besten Fokussierungsposition zwar sehr schnell, die Breite der Spitze ist aber zu eng. Deshalb ist der Arbeitsbereich der Kamera sehr klein.

Aus der o. g. Diskussion kann man zusammenfassen:

- a. SMD ist die beste FF. VAR, TV und LO sind auch günstige Fokussierungsfunktionen.
- b. LS und SL' sind zu langsam zur Fokussierung. Außerdem ist ihre Genauigkeit schlecht.
- c. Die Fokussierungsfunktion ENT ist zwar schnell, aber ihre Genauigkeit ist zu schlecht.
- d. Das Differential der Varianz als Fokussierungsfunktion weist gegenüber der Varianz keine Vorteile auf.

5. Fokussierung von Bildausschnitten

5.1 Grundidee der Fokussierung von Bildausschnitten

Die Szene besteht normalerweise aus verschiedenen Objekten, die sich nicht immer in einer Ebene befinden. Deshalb sind die Objektweiten dieser Objekte unterschiedlich. Sogar die Oberfläche eines Objektes ist nicht flach. Wenn wir die Oberfläche des Objektes in verschiedene Regionen aufteilen, sind die Abstände zwischen der Linse der Kamera und solchen Regionen unterschiedlich. Solange sich alle Objekte in der Szene nicht im Schärfentiefbereich der Kamera befinden, kann die Kamera solche Objekte nicht gleichzeitig im Fokus haben, z.B. wenn man zwei Gebäude fotografiert, die verschieden weit von der Kamera entfernt sind. Wenn die Kamera so eingestellt wird, daß das Bild eines Gebäudes scharf ist, ist das des anderen vielleicht nicht immer scharf.

Die oben erwähnten Fokussierungsfunktionen sind zur Fokussierung des Objektes mit flacher Oberfläche entworfen. Um die Genauigkeit der Fokussierung zu garantieren, werden die benutzten Bildpunkte sich im ganzen Bild verteilen. Deshalb drücken die Fokussierungsfunktionen den Fokussierungszustand des ganzen Bildes aus. Wenn die Szene verschieden weit von der Linse der Kamera entfernt ist, werden die Fokussierungsfunktionen mehrere Maxima in verschiedenen Positionen haben.

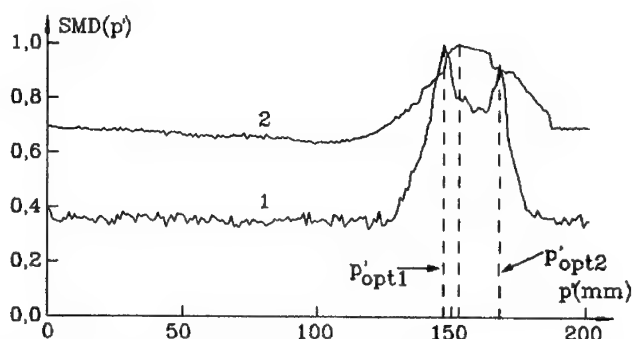


Bild 5.1: SMD eines Objektes mit Tiefsprung.

Bild 5.1 zeigt die SMD eines Objektes mit einem Tiefensprung. Die Oberfläche dieses Objektes besteht aus zwei Flächen. Die Objektweiten von beiden Flächen sind unterschiedlich. Deshalb treten bei der SMD zwei lokale Extremwerte in den Positionen p'_{opt1} und p'_{opt2} in der Kurve 1 auf. Die Kamera kann aber nur einmal eine Fläche fokussieren. Normalerweise wird mit Hilfe der Fokussierungsfunktion die Position des ersten Extremwertes als die beste Fokussierungsposition gefunden. Solange sich die beiden Flächen nicht im Schärfentiefenbereich der Kamera befinden, hat die Fokussierungsfunktion zwei lokale Extremwerte (Kurve 1 in Bild 5.1). Durch die Positionen der beiden Extremwerte kann man den Abstand der beiden Flächen bestimmen. In Kurve 1 ist $p'_{opt1} = 147$ mm und $p'_{opt2} = 168$ mm, deshalb sind die beiden Flächen 21 mm voneinander entfernt. Wenn beide Flächen im Schärfentiefenbereich der Kamera liegen, tritt bei der Fokussierungsfunktion nur ein Maximum auf, aber die Kurve der FF in der Nähe der besten Fokussierungsposition ist sehr flach (Kurve 2 in Bild 5.1).

Durch die o. g. Analyse kann man erkennen: wenn die benutzten Bildpunkte sich im ganzen Bild verteilen, ist die Fokussierung der Szene mit verschiedenen Objektweiten unmöglich. Um irgendwo das Objekt zu fokussieren, muß zuerst die zu fokussierende Position ausgewählt werden. Die Berechnung der FF benutzt nur die Bildpunkte in der Umgebung dieser Position. Es werden so wenig wie möglich Bildpunkte benutzt, damit die Fokussierungsfunktion weniger von anderen Flächen beeinflusst wird.

Um die zu fokussierende Position des Objektes zu bestimmen, muß das Objekt zuerst grob fokussiert werden. Durch die grobe Fokussierung bekommt man ein unscharfes Bild des Objektes. Mit Hilfe dieses unscharfen Bildes wird die zu fokussierende Position ausgewählt und deren Koordinaten (x,y) eingegeben. In Gauß'scher Näherung [17] soll jedem Bildpunkt die entsprechende Position vom Objekt zugeordnet werden. Nach der Eingabe der Koordinaten (x,y) berechnet der Computer die Fokussierungsfunktion mit den Grauwerten des quadratischen Umgebungsfensters von (x,y). Wegen des guten Charaktes wird die SMD zur Fokussierung der relativen Position benutzt.

Nach der Wahl der Fokussierungsfunktion ist es wichtig, die Größe der benutzten Bildpunkte festzulegen. Sind es zu wenig Bildpunkte, wird die Fokussierungsfunktion durch Rauschen stark beeinflusst. Die Größe der benutzten Bildpunkte wird durch folgendes Experiment bestimmt.

Im Experiment werden die Kurven der SMD für verschiedene Größen der benutzten Bildpunkte gemessen. Als Größen werden folgende vier Fälle betrachtet:

- a) $16 \times 16 = 256$ (Bildpunkte)
- b) $32 \times 32 = 1024$ (Bildpunkte)
- c) $64 \times 64 = 4096$ (Bildpunkte)
- d) $128 \times 128 = 16384$ (Bildpunkte)

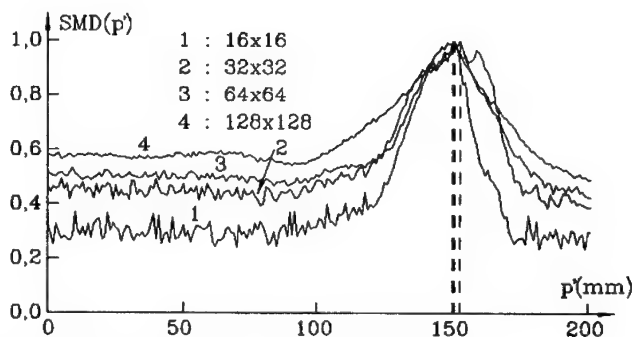


Bild 5.2: SMD für verschiedene Anzahl von Bildpunkten.

Die Kurve 1 in Bild 5.2 zeigt die SMD bei der Größe 16×16 . Weil die Berechnung der SMD zu wenige Bildpunkte benutzt hat, tritt die Kurve in der besten Fokussierungsposition zwar als Maximum auf, sie hat aber große Schwingungen im unwirksamen Bereich der Kamera. Diese Schwingungen können den Fokussierungsprozeß unterbrechen. Wenn die Größe der benutzten Bildpunkte 128×128 ist, kann der Einfluß des Rauschens vernachlässigt werden (Kurve 4), aber für die Fokussierung einer relativen Position sind 128×128 Bildpunkte zu viel.

Kurve 2 und Kurve 3 in Bild 5.2 entsprechen der SMD von 32×32 und 64×64 Bildpunkten. In beiden Kurven sind die Einflüsse des Rauschens im Vergleich zur Kurve 1 schwach. Deshalb wird als Größe der Bildpunkte 32×32 gewählt. Durch die Vorverarbeitung der Bilder kann damit der Einfluß des Rauschens vermieden oder zumindest verringert werden.

5.2 Vorverarbeitung zur Reduzierung des Rauscheinflusses

Bei den meisten Diskussionen über Fokussierungsalgorithmen wird angenommen, daß es möglich ist, die exakte Beziehung zwischen dem Objekt und dessen Bild zu kennen [18]. Unglücklicherweise sind alle Bilder aber immer durch das Rauschen verfälscht. Bilder werden gewöhnlich während ihrer Erzeugung durch das abbildende System durch zufällige Störungen, d.h. Rauschen, beeinflußt. Der Beitrag des Rauschens kommt aus folgenden Quellen: das Schrotrauschen im Sensor, das thermische Rauschen in den Verstärkungsbestandteilen.

Den realen Grund zur Bildvorverarbeitung vor der eigentlichen Berechnung der Fokussierungsfunktion kann man durch die Fokussierungsstrategie deutlich erkennen. Alle Fokussierungskriterien haben nämlich eine gemeinsame Eigenschaft, sie sind abhängig von den Fokussierungsfunktionen, die ein Maximum oder Minimum in der besten Fokussierungsposition haben. Die Fokussierungsstrategie ist also so, daß der Computer die Fokussierungsfunktion berechnet und die Position der Kamera verändert. Dann berechnet der Computer die FF nochmal und vergleicht die FF in beiden Positionen. Wenn die FF abnimmt, fährt die Kamera zurück, sonst fährt die Kamera vor und zwar soweit, bis die beste Fokussierungsposition gefunden ist.

Es wird angenommen, daß wegen des Rauscheinflusses die FF ein lokales Maximum vor der besten Fokussierungsposition hat. In diesem Fall wird diese Position als beste Fokussierungsposition gefunden.

Um effektive Algorithmen für die Bildvorverarbeitung auslegen zu können, muß man die Ursachen des Bildrauschens und ihre Auswirkungen auf den vollständig digitalisierten Sensorausgang kennen. Die Bildvorverarbeitung besteht im wesentlichen aus einem Abschätzungsprozeß, mit dem ein "ideales" Bild erzeugt werden soll, aus dem alle Einflüsse des Rauschens entfernt worden sind (Bild 5.3) und basiert darauf, ein rekonstruiertes Bild $g'(x,y)$ zu erzeugen, das eng mit dem unverrauschen Bild $g_0(x,y)$ in Beziehung steht, das der Sensor erzeugt hat.

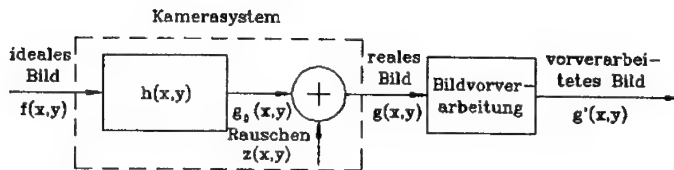


Bild 5.3: Blockschaltbild eines typischen Bildsystems.

Es gibt zwei Haupttypen des Rauschens:

a) Additives Rauschen.

Das Modell des additiven Rauschens verwendet man dann, wenn das Bild am Ausgang des Systems oder nach einem beliebigen Transformationsschritt als Summe des unverrauschten Bildes und eines zufälligen Rauschens betrachtet werden kann. Damit kann beispielsweise die Wirkung des Schrotrauschens im Sensor und des thermischen Rauschens in den Verstärkungsbestandteilen beschrieben werden.

Wenn sich die Wirkung des Rauschens nicht über das ganze Bild erstreckt, sondern nur an zufällig verteilten Bildpunkten existiert, an denen sich der Grauwert um eine zufällige Größe ändert, nennt man dies Impulsrauschen. Impulsrauschen ist eine Art von additivem Rauschen und ist charakteristisch für Systeme zur Bildübertragung mittels nichtlinearer Modulationsarten (Frequenzmodulation, Pulsdauermodulation usw.) auf Funkkanälen und digitalen Systemen zur Bildübertragung und -speicherung.

Additives Rauschen beeinflusst den Grauwert jedes Bildpunktes unabhängig voneinander (d.h. es verändert die Graustufe des individuellen Punktes).

b) Multiplikatives Rauschen.

Das multiplikative Modell wird dann angewandt, wenn man annehmen kann, daß das Bild mit dem zufälligen Rauschen multipliziert wird. Das ist z.B. in erster Näherung der Fall beim Rauschen von Sekundärelektronenvervielfachern und beim Diffusionsrauschen in kohärent-optischen und holographischen Abbildungssystemen. Das multiplikative Rauschen beeinflusst die Grauwerte mehrerer Bildpunkte

gleichzeitig. In der vorliegenden Arbeit steht das additive Rauschen im Mittelpunkt der Betrachtungen.

Zur Berücksichtigung dieses Rauschens ist es notwendig, die statistischen Charakteristika des Rauschens zu kennen. Manchmal kann man diese Charakteristika bestimmen, wenn man von der Struktur und den Charakteristika des optischen Systems ausgeht. Beim Fehlen solcher Daten ist eine Abschätzung der Rauschcharakteristik aus schon erzeugten Bildern oder aus einem Satz von Bildern mit homogenem Rauschen notwendig. In solchen Fällen müssen die statistischen Charakteristika des Rauschens aus der Messung der statistischen Charakteristika des beobachteten Bildes abgeleitet werden.

Um die Rauschparameter im beobachteten Gemisch zu schätzen, ist es notwendig, Signal und Rauschen im Bild zu trennen. Dies ist wiederum nur möglich, wenn man die statistischen Parameter vom Signal und Rauschen kennt.

Glücklicherweise ist jedoch in der Mehrzahl der praktischen Fälle das Rauschen, das im Bild zu unterdrücken ist, im statistischen Sinn sehr einfach. Das Rauschen kann durch eine geringe Zahl von Parametern beschrieben werden. Daher läßt sich diese vereinfachte Aufgabe der Parameterschätzung des Rauschens selbst bei einer sehr groben und nicht sehr detaillierten a priori Vorgabe der statistischen Parameter des gemessenen Bildes mit relativ einfachen Mitteln lösen. Es ist lediglich notwendig, aus allen der Messung zugänglichen statistischen Signalparametern diejenigen auszuwählen, in denen die Störung des Bildes durch das Rauschen in möglichst einfach nachweisbarem anormalem Verhalten zum Ausdruck kommt.

Weil die Analysierung der Auswirkung des Rauschens in vielen Fällen möglich ist, kann man die Auswirkung des Rauschens verringern.

1) Räumliche Glättung.

Die meisten Bildvorverarbeitungstechniken basieren auf digitaler Filterung, d.h. auf der Eliminierung unerwünschter Effekte aus dem "realen" Bild, das der Sensor und das optische System liefern [19]. Die Glättung der Grauwerte eines Bildes ist eine Art der digitalen Filterung. Die einfachste Form der Glättung der Grauwerte eines Bildes ist die Berechnung des Mittelwertes einer vorgegebenen Umgebung eines Bildpunktes. Aus Gründen der Symmetrie wird meistens eine quadratische Umgebung verwendet. Die Größe des quadratischen Umgebungsfensters wird mit $m \times m$ bezeichnet. Der Operator

$$g'(x,y) = \frac{1}{m} \sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k g(x-\tilde{x}, y-\tilde{y}) \quad (5.1)$$

mit $m = 3, 5, 7, \dots$ und $k = (m-1)/2$ weist dem Bildpunkt $g'(x,y)$ den Mittelwert der $m \times m$ -Umgebung des Bildpunktes $g(x,y)$ als neuen Grauwert zu. Für die meisten Fälle ist $m = 3$, dann werden folgende Bildpunkte in die Berechnung einbezogen:

$$\begin{array}{lll} g(x-1,y-1) & , & g(x-1,y) & , & g(x-1,y+1) & , \\ g(x,y-1) & , & g(x,y) & , & g(x,y+1) & , \\ g(x+1,y-1) & , & g(x+1,y) & , & g(x+1,y+1) & . \end{array}$$

Bilder, die mit diesem Operator, der auch bewegter Mittelwert genannt wird, verarbeitet werden, wirken im Vergleich zum Original weicher oder etwas unschärfer. In Bildbereichen mit homogenen Grauwerten hat der bewegte Mittelwert keine Auswirkung, wogegen Grauwertübergänge geglättet werden. Durch die räumliche Filterung (5.1) wird das dem originalen Bild überlagerte Rauschen verringert. Dies läßt sich rechnerisch nachweisen. Es wird angenommen, daß dem originalen Bild g ein Rauschanteil additiv überlagert ist:

$$g(x,y) = g_0(x,y) + z(x,y) \quad (5.2)$$

Der Rauschanteil $z(x,y)$ wird durch Zufallsvariable $Z_{(x,y)}$ beschrieben, die den Erwartungswert 0 und die Streuung σ^2 haben. Die Streuung des Rauschens wird auch als Rauschleistung bezeichnet. Die Zufallsvariablen seien unabhängig und

unkorreliert. Damit läßt sich für $g'(x,y)$ schreiben:

$$g'(x,y) = \frac{1}{m^2} \cdot \left[\sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k g_0(x-\tilde{x}, y-\tilde{y}) + \sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k z(x-\tilde{x}, y-\tilde{y}) \right] \quad (5.3)$$

Für den Erwartungswert und die Streuung der Zufallsvariablen

$$Z = \frac{1}{m^2} \cdot \sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k Z_{(x-\tilde{x}, y-\tilde{y})}$$

ergibt sich:

$$\begin{aligned} E(Z) &= E \left[\frac{1}{m^2} \cdot \sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k Z_{(x-\tilde{x}, y-\tilde{y})} \right] = \\ &= \frac{1}{m^2} \cdot \sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k E \left[Z_{(x-\tilde{x}, y-\tilde{y})} \right] = 0 \quad , \end{aligned} \quad (5.4)$$

$$\begin{aligned} E(Z^2) &= E \left[\left[\frac{1}{m^2} \cdot \sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k Z_{(x-\tilde{x}, y-\tilde{y})} \right]^2 \right] = \\ &= \frac{1}{m^4} \cdot E \left[\left[\sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k Z_{(x-\tilde{x}, y-\tilde{y})} \right]^2 \right] = \\ &= \frac{1}{m^4} \cdot \sum_{\tilde{x}=-k}^k \sum_{\tilde{y}=-k}^k E \left[\left[Z_{(x-\tilde{x}, y-\tilde{y})} \right]^2 \right] = \frac{1}{m^2} \cdot \sigma^2 \quad . \end{aligned} \quad (5.5)$$

Die Beziehungen (5.4) und (5.5) bestätigen die obigen Vermutungen: Der Erwartungswert bleibt erhalten und die Streuung des Rauschanteils wird um den Faktor $1/m^2$ reduziert.

Der gleitende Mittelwert kann auch als Faltungsoperation aufgefaßt werden. Da beim gleitenden Mittelwert nur über eine endliche $m \times m$ -Umgebung des Bildes gemittelt wird, kann (5.2) als Faltung des Bildausschnittes mit einer Maske

$\underline{h} = [h(\tilde{x}, \tilde{y})]$ geschrieben werden:

$$g'(x,y) = \frac{1}{m^2} \cdot \sum_{\tilde{x}=0}^{m-1} \sum_{\tilde{y}=0}^{m-1} g(x+k-\tilde{x}, y+k-\tilde{y}) \cdot h(\tilde{x}, \tilde{y}) \quad . \quad (5.6)$$

Die Maske H hat beim gleitenden Mittelwert mit $m = 3$ das Aussehen

$$H = [h(\tilde{x}, \tilde{y})] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} .$$

Die zwei anderen Masken, die häufig zur räumlichen Glättung benutzt werden, haben das Aussehen

$$H_1 = \begin{bmatrix} 0 & 2.25 & 0 \\ 2.25 & 0 & 2.25 \\ 0 & 2.25 & 0 \end{bmatrix} ,$$

$$H_2 = \begin{bmatrix} 1.125 & 1.125 & 1.125 \\ 1.125 & 0 & 1.125 \\ 1.125 & 1.125 & 1.125 \end{bmatrix} .$$

Setzt man H_1 und H_2 in Gleichung (5.6) ein, ergibt sich

$$g'(x, y) = \frac{g(x-1, y) + g(x, y-1) + g(x, y+1) + g(x+1, y)}{4} \quad (5.7)$$

und

$$g'(x, y) = \frac{g(x-1, y-1) + g(x-1, y) + g(x-1, y+1) + g(x, y-1) + g(x, y+1) + g(x+1, y-1) + g(x+1, y) + g(x+1, y+1)}{8} . \quad (5.8)$$

Die Auswirkung der räumlichen Glättung kann einfach mit Hilfe des Laplace-Operators erklärt werden. Der Grauwert eines Bildes ist die Funktion von zwei Ortsvariablen (x, y) . Es ist angenommen, daß diese Funktion eine Oberfläche im dreidimensionalen Raum ist. Zur Reduzierung des Rauschens braucht man nur die Krümmung der Oberfläche in jedem Punkt zu verringern. Am besten ist es, wenn diese Krümmung Null ist. Das wird genau dann passieren, wenn der Laplace-Operator gleich Null ist, also

$$\nabla^2 g(x, y) = \frac{\partial^2 g(x, y)}{\partial x^2} + \frac{\partial^2 g(x, y)}{\partial y^2} = 0 . \quad (5.9)$$

Die Approximation der zweiten Ableitung lautet:

$$\frac{\partial^2 g(x,y)}{\partial x^2} = g(x+1,y) - 2 \cdot g(x,y) + g(x-1,y) \quad ,$$

$$\frac{\partial^2 g(x,y)}{\partial y^2} = g(x,y+1) - 2 \cdot g(x,y) + g(x,y-1) \quad .$$

Setzt man $\partial^2 g(x,y)/\partial x^2$ und $\partial^2 g(x,y)/\partial y^2$ in (5.9) ein, ergibt sich die Gleichung (5.7).

Der Hauptnachteil der räumlichen Glättung ist ihre Tendenz zur Trübung des Bildes. Dies ist besonders schlimm für den Fokussierungsprozeß. Alle oben erwähnten Glättungsalgorithmen funktionieren nämlich als Tiefpaßfilter.

Die Auswirkung der Glättung erkennt man in Bild 5.4. Es zeigt die Fokussierungsfunktion vor und nach der räumlichen Glättung.

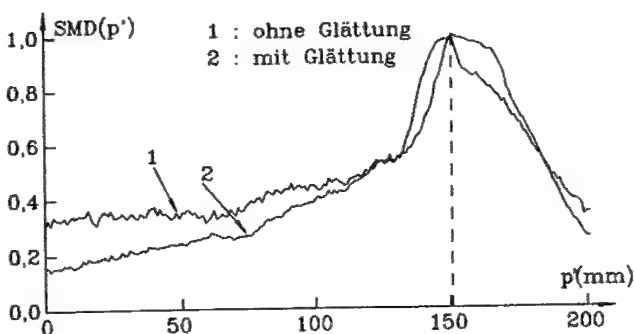


Bild 5.4: SMD vor und nach der räumlichen Glättung.

Kurve 1 zeigt die SMD des originalen Bildes, das nicht durch räumliche Glättung behandelt ist. Deswegen hat sie große Schwingungen im unwirksamen Bereich der Kamera. In Kurve 2 ist die SMD nach der räumlichen Glättung mit (5.4) gezeigt. Die Schwingungen in Kurve 2 sind deutlich verringert, aber die Kurve 2 ist im Vergleich zur Kurve 1 flach, weil durch die Glättungsoperation das originale Bild verschmiert wird.

2) Zeitliche Glättung

Die räumliche Glättung kann zwar den Rauschanteil im Bild verringern, hat jedoch zu verwaschenen Kanten geführt. Die zeitliche Glättung kann diesen Nachteil vermeiden. Für die zeitliche Glättung wird ein Zeitreihenbild

$\underline{g} = [g(x,y,t)]$, $t = 0, 1, \dots, T-1$, vorausgesetzt. Es kann sich dabei z.B. um eine Bildsequenz handeln, die von einer CCD-Kamera erzeugt wird, die fest installiert ist und einen definierten Bildausschnitt liefert. Ist der Ausschnitt nicht bewegt und sind die Parameter der Kamera unverändert, so sind die Zeitreihenbilder im wesentlichen alle gleich, bis auf das Rauschen, das durch die Aufnahmebedingungen in die Grauwerte des Bildes einfließt. Eine Glättung der zeitlich aufeinanderfolgenden Bilder liefert dann ein Bild, in dem die Rauscheinflüsse geringer sind (auch Gauß'sches Rauschen mit Erwartungswert 0):

$$g'(x,y) = \frac{1}{T} \cdot \sum_{t=0}^{T-1} g(x,y,t) \quad . \quad (5.10)$$

Das geglättete Bild \underline{g}' ist in der Bildqualität verbessert, wobei die Glättung in (5.10) die Schärfe des Bildes nicht beeinflusst hat.

Durch die zeitliche Glättung (5.10) wird der Grauwert $g'(x,y)$ in jedem Bildpunkt durch die geglätteten Grauwerte $g(x,y,t)$ im gleichen Bildpunkt zu verschiedenen Zeitpunkten ersetzt. Je länger die zeitliche Glättung dauert, desto besser ist die Reduzierung des Rauschanteils im Bild. Durch die zeitliche Glättung kann die Streuung des Rauschens um den Faktor $1/T$ reduziert werden. Der Grauwert $g(x,y,t)$ wird wieder in der Form

$$g(x,y,t) = g_0(x,y,t) + z(x,y,t) \quad (5.11)$$

angesetzt. Wiederum werden die Rauschanteile $z(x,y,t)$ durch Zufallsvariable $Z(x,y,t)$ beschrieben. Dann läßt sich für $g'(x,y)$ schreiben:

$$g'(x,y) = \frac{1}{T} \cdot \left[\sum_{t=0}^{T-1} g_0(x,y,t) + \sum_{t=0}^{T-1} z(x,y,t) \right] \quad . \quad (5.12)$$

Für die Streuung der Zufallsvariablen

$$Z = \frac{1}{T} \cdot \sum_{t=0}^{T-1} Z(x, y, t)$$

ergibt sich

$$\begin{aligned} E(Z^2) &= E \left[\left[\frac{1}{T} \cdot \sum_{t=0}^{T-1} Z(x, y, t) \right]^2 \right] = \\ &= \frac{1}{T^2} \cdot E \left[Z_{(x, y, 0)}^2 + Z_{(x, y, 1)}^2 + \dots + Z_{(x, y, T-1)}^2 \right] = \\ &= \frac{1}{T^2} \cdot (T \cdot \sigma^2) = \frac{1}{T} \cdot \sigma^2 \end{aligned} \quad (5.13)$$

Sofern das Bild stationär ist, ist die zeitliche Glättung eine bessere Methode zur Reduzierung des Rauschens als die räumliche Glättung.

Die Auswirkung der zeitlichen Glättung auf die Fokussierungsfunktion ist in Bild 5.5 demonstriert.

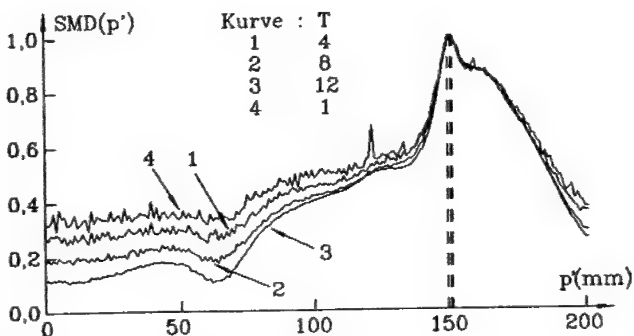


Bild 5.5: SMD nach der zeitlichen Glättung.

Kurve 1, 2 und 3 entsprechen der SMD nach der zeitlichen Glättung von $T = 4, 8$ und 12 . Durch die zeitliche Glättung sind die Schwingungen der Fokussierungsfunktion stark verringert. Die Kurve der SMD ist im Vergleich zur SMD des originalen Bildes nicht verflacht.

5.3 Kompensierung der Zoom-Auswirkung

Parallel verlaufende Linien scheinen zusammenzulaufen; dieses Phänomen läßt sich gut an Bahngleisen beobachten. Stellt man sich vor, in Bild 5.6 wären A und B Bahnschwellen, die parallele Schienen trennen [20],[21].

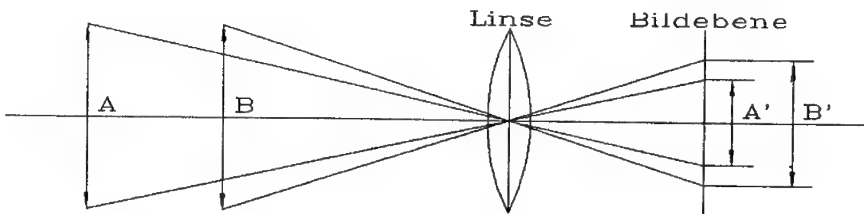


Bild 5.6: Schematische Darstellung der Zoom-Auswirkung.

Die Schwellen A und B sind gleich groß. Das Bild der entfernteren Schwelle A ist kleiner als B, deshalb müssen die Schienen, die auf diesen Schwellen liegen, bei A näher beieinander sein als bei B. Die parallelen Eisenbahnschienen scheinen in beiden Richtungen zusammenzulaufen. Das ist die typische Zoom-Auswirkung.

Im alltäglichen Leben haben wir alle schon solche Erfahrungen gemacht [22]: Bewegt man sich durch eine Welt von stationären Objekten, so ist die auf der Netzhaut projizierte visuelle Welt scheinbar vorbeigeflogen. Tatsächlich scheint die Welt für eine bestimmte Richtung der Schubbewegung und eine bestimmte Blickrichtung aus einem besonderen Fluchtpunkt, also dem sogenannten "Focus Of Expansion" (FOE), herauszufließen. Jede Richtung der Bewegung und des Blickes verursacht eine einzige FOE. Für die automatische Fokussierung in der vorliegenden Arbeit ist die Bewegungsrichtung der Kamera senkrecht zur Szene. Deshalb befindet sich die FOE in der Mitte des Bildes.

Wenn die Kamera die relative Position des Objektes fokussiert, muß zuerst die zu fokussierende Position eingegeben werden. Solange sich diese Position nicht in der FOE, also der Mitte des Bildes, befindet, wird sich das Bild von dieser Position aus während des Fokussierungsprozesses verändern.

Die Bewegung der Kamera führt die Bewegung im Bild. Die Bewegung von einem Bild kann durch ein zweidimensionales Verschiebungsvektorfeld beschrieben werden. Jeder Vektor dieses Feldes verbindet einander entsprechende Objektpunkte. Für das Verschiebungsvektorfeld bzw. einen einzelnen Verschiebungsvektor werden im folgenden die Abkürzungen VVF bzw. VV benutzt.

Die vollständige Bestimmung des VVF ist nur möglich, wenn die Bildpunkte in dem einen Bild eindeutigen Punkten im zweiten Bild zugeordnet werden können. Wenn das Objekt starr ist und die Bewegung von einem Bild durch die Bewegung der Kamera senkrecht zum Objekt verursacht ist, ist die vollständige Bestimmung des VVF garantiert. Das ist genau die Situation der Fokussierung zur relativen Position des Objektes. In der Berechnung der Fokussierungsfunktion der relativen Position muß die Bewegung des Bildes, als die sogenannte Zoom-Auswirkung kompensiert werden. Weil das Objekt unbeweglich ist und nur die Kamera sich während des Fokussierungsprozesses bewegt, kann die Zoom-Auswirkung leicht kompensiert werden.

Vor der Kompensation der Zoom-Auswirkung muß das VVF überprüft werden. Wie sieht das VVF in der Bildebene aus? Wenn wir es vollständig kennen, läßt sich damit eindeutig auf die Bewegungsform im dreidimensionalen Raum schließen. Bild 5.7 stellt für das VVF eine quadratische Fläche dar. Die Bewegung ist mit einem charakteristischen VVF verknüpft. Alle VV, die durch die Bewegung der Kamera senkrecht zum Objekt (weg vom Objekt) verursacht sind, sind auf die FOE gerichtet.

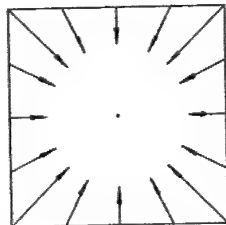


Bild 5.7: Schematische Darstellung eines VVF.

Die FOE befindet sich genau in der Mitte des Bildes. Nachfolgend werden die Bewegungsvorgänge näher untersucht, dabei wird nur die obenerwähnte Bewegungsart betrachtet, die Kamera bewegt sich also senkrecht zum Objekt.

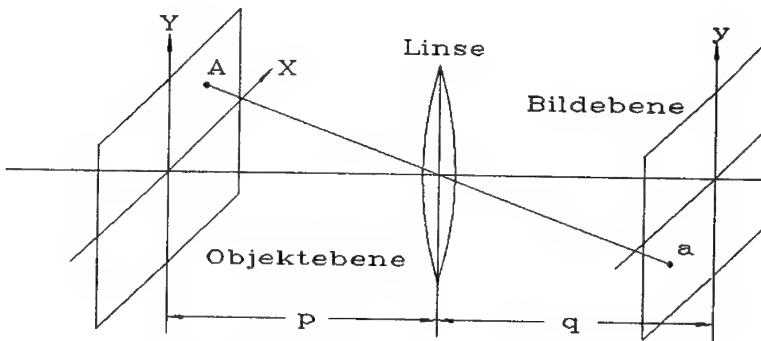


Bild 5.8: Objekt- und Bildebene.

Während des Fokussierungsprozesses ist die Bildweite q unverändert. Wenn die Kamera sich bewegt, wird die Objektweite p ständig verändert. Die Objekzebene (X, Y) wird durch das optische System auf die Bildebene (x, y) abgebildet. In Gauß'scher Näherung soll jedem Punkt in der Objekzebene das entsprechende Bild in der Bildebene zugeordnet sein (Bild 5.8).

Der Übergang vom Objekt- zum Bildkoordinatensystem wird durch eine Koordinatentransformation beschrieben.

Wenn die Koordinaten von der zu fokussierenden Position in der Bildebene (x_1, y_1) sind, kann die entsprechende Position in der Objekzebene (X_1, Y_1) durch die geometrische Optik berechnet werden. Es ist

$$\begin{cases} X_1 = -\frac{x_1}{q} \cdot p_1 \\ Y_1 = -\frac{y_1}{q} \cdot p_1 \end{cases} \quad (5.14)$$

Jetzt ist die zu fokussierende Position (X_1, Y_1) in der Objektebene bestimmt. Während des Fokussierungsprozesses verändert sich die Objektweite p . Deshalb wird das Bild von (X_1, Y_1) auch verändert und zwar wie folgt:

$$\begin{cases} x = -\frac{X_1}{p} \cdot q = x_1 \cdot \frac{p_1}{p} , \\ y = -\frac{Y_1}{p} \cdot q = y_1 \cdot \frac{p_1}{p} . \end{cases} \quad (5.15)$$

Um die Koordinaten (X_1, Y_1) genau zu berechnen, muß die Objektweite p_1 bekannt sein. Aber das ist unmöglich, sonst bräuchte man das Objekt nicht zu fokussieren. Deshalb muß das Objekt zuerst grob fokussiert werden. Durch die grobe Fokussierung kann die Objektweite p_1 annähernd bestimmt werden, d.h. die Zoom-Auswirkung kann annähernd kompensiert werden. Folgend wird der Prozeß der Fokussierung zur relativen Position zusammengefaßt:

- 1) Das Objekt wird zuerst grob fokussiert. Danach wird die zu fokussierende Position (x_1, y_1) in der Bildebene ausgewählt und die Objektweite p_1 bestimmt.
- 2) Die zu (x_1, y_1) entsprechenden Koordinaten (X_1, Y_1) werden nach (5.14) berechnet. Während des ganzen Fokussierungsprozesses werden (X_1, Y_1) festbleiben.
- 3) Die Kamera wird in der Nähe von p_1 die beste Fokussierungsposition von (X_1, Y_1) suchen. Vor der Berechnung der FF wird die Zoom-Auswirkung kompensiert. Nach der Bewegung der Kamera werden die zu (X_1, Y_1) entsprechenden Koordinaten (x, y) durch (5.15) ermittelt. Anschließend wird die FF mit den Umgebungsbildpunkten von (x, y) berechnet.

6. Fokussierung von beweglichen Objekten

6.1 Das Prinzip der Fokussierung von beweglichen Objekten

Die bisherigen Betrachtungen beschäftigen sich nur mit der Fokussierung von unbeweglichen Objekten. Für ein derartiges Fokussierungsproblem spielt die Fokussierungsfunktion die entscheidende Rolle. Mit Hilfe der sogenannten "Hill-Climbing" Strategie [23], [24] wird die Position, wo das Maximum oder Minimum der Fokussierungsfunktionen auftritt, gesucht. Es ist sehr wichtig, daß sich während des Fokussierungsprozesses die zu fokussierenden Objekte nicht bewegen.

Es gibt zwei Bewegungsarten vom Objekt relativ zur Kameralinse. Erstens bewegt sich das Objekt querab zur Kameralinse, d.h. der Abstand zwischen der Kameralinse und dem Objekt bleibt näherungsweise unveränderlich, der von der Kamera aufgenommene Objektabschnitt ist aber unterschiedlich. Für viele Objekte ist die Textur der Oberfläche von Ort zu Ort unterschiedlich. Wenn dies der Fall ist, ist die Fokussierungsfunktion nicht mehr unimodular. Tritt auf der ganzen Oberfläche des Objektes nur eine einzige Textur auf, ist eine Fokussierung eines solchen beweglichen Objektes mit Hilfe der Fokussierungsfunktion allerdings auch bei Bewegung möglich.

Bei der anderen Bewegungsart bewegt sich das Objekt senkrecht zur Kameralinse. Diese Bewegungsart verändert den Abstand zwischen Kameralinse und dem Objekt. Die Veränderung dieses Abstandes ändert die Verschiebung der besten Fokussierungsposition. Es ist klar, daß die Fokussierungsfunktionen für solche beweglichen Objekte ungeeignet sind.

In vielen Fällen bewegt sich das Objekt sowohl querab als auch senkrecht zur Kameralinse. Die Fokussierungsfunktion ist in diesen Fällen auch ungeeignet zur Fokussierung von beweglichen Objekten.

Es gibt heute viele Kameras, die die Fähigkeit zur automatischen Fokussierung von beweglichen Objekten haben. Das Fokussierungsprinzip einer solchen Kamera ist in zwei grundlegende Kategorien eingeteilt, nämlich als aktives System oder als passives System [25], [26]. Beide Systeme brauchen ein zusätzliches Gerät. Bei der aktiven automatischen Fokussierung wird elektromagnetische oder Ultraschallstrahlung ausgesendet. Erreicht diese Strahlung das Ziel, wird ein

Teil der Strahlung reflektiert. Die reflektierte Strahlung wird durch einen angepaßten Empfänger registriert. Gemäß der Laufzeit der Strahlung kann der Abstand zwischen dem Objekt und der Kamera bestimmt werden. Anschließend ist die Kameralinse richtig eingestellt.

Die erste Kamera Pentax MEF mit passiver Methode zur Realisierung der automatischen Fokussierung ist im Jahr 1981 entwickelt worden [27]. Das automatische Fokussierungssystem von Pentax ist das sogenannte EFC (Electronic-Focus Control) System. Statt der Untersuchung der Fokussierungszustände in verschiedenen Fokussierungspositionen wird im EFC-System nur das Erscheinen des Bildes auf beiden Seiten der Bildebene untersucht (Bild 6.1).

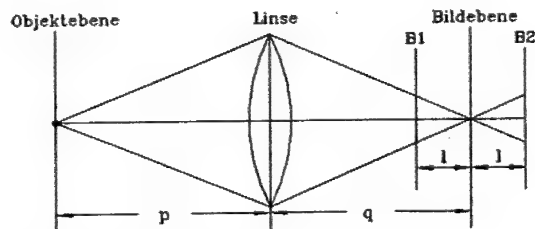


Bild 6.1: Prinzip des passiven Systems.

Dazu werden zwei Detektoren B1, B2 in gleichem Abstand l von beiden Seiten der Bildebene eingestellt. Wenn die Kamera sich in der besten Fokussierungsposition befindet, sind die elektrischen Signale von beiden Detektoren identisch. Wenn das Bild in der Bildebene nicht im Fokus ist, ist das auf beide Detektoren einfallende Licht unterschiedlich. Deshalb gibt es einen Unterschied zwischen beiden elektrischen Signalen. Durch diesen Unterschied kann die Kamera wieder die Szene fokussieren. Das Vorzeichen dieses Unterschieds bestimmt die Richtung der Linsenbewegung. Solange dieses Vorzeichen sich nicht verändert, wird die Linse in einer Richtung weiter verschoben, bis der Unterschied verschwindet.

In der vorliegenden Arbeit wird das Prinzip der Fokussierung eines beweglichen Objektes anders als bei der oben erwähnten Methode gewählt. Im System wird ein zusätzliches Lasergerät benutzt, aber ohne Empfänger zur Registrierung der reflektierten Laserstrahlung. Die Laserstrahlung zeigt hohe zeitliche Kohärenz

(lange zusammenhängende Wellenzüge) und hohe räumliche Kohärenz. Sie ist eine typisch punktförmige Lichtquelle.

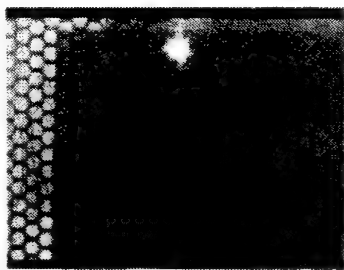
Wegen der hohen räumlichen Kohärenz ist die Laserstrahlung auf einem kleinen Gebiet des Objektes konzentriert. Die Beleuchtungsstärke von diesem Gebiet ist viel höher als die des übrigen Gebietes vom Objekt. Deshalb sind die Grauwerte in diesem Gebiet erheblich größer als die von anderen Bildpunkten. Mit Hilfe dieser Tatsache ist es leicht, die Koordinaten des Bildes vom Laserfleck zu ermitteln.

Die Laserstrahlung ist in einem bestimmten Winkel zur optischen Achse der Kamera ausgerichtet. Damit kann der Laserfleck in verschiedenen Positionen an der Oberfläche des Objektes einfallen, wenn sich die Position des Objektes verändert. Wenn die anderen Parameter der Kamera festbleiben, kann der Abstand zwischen der Kamera und dem Objekt mit Hilfe der Bildkoordinaten des Laserflecks direkt berechnet werden.

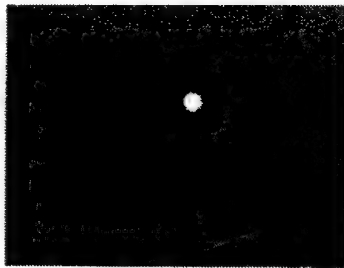
Im System wird ein HeNe-Lasergerät benutzt. Die Wellenlänge dieses Lasergerätes ist 632,8 nm. Deshalb ist die Laserstrahlung rot. Unter bestimmter Bedingungen, nämlich daß die Parameter der Kamera unverändert bleiben, können über die Koordinaten des Laserflecks im Bild der Abstand zwischen der Kameralinse und dem Objekt vollständig bestimmt werden. Bild 6.2 zeigt die Bilder des durch eine Laserstrahlung gerichteten Objektes in verschiedenen Objektweiten.

In Bild 6.2a sind die analogen Signale des Bildes vor der Digitalisierung nicht vorverarbeitet. Die Grauwerte des Laserflecks sind nicht so deutlich von der Umgebung getrennt. Um die Grauwerte des Laserflecks problemlos von der Umgebung zu unterscheiden, muß das analoge Signal der CCD-Kamera vor der Digitalisierung behandelt werden. Die folgende Gleichung stellt die Beziehung zwischen der eingegebenen Spannung $v(x,y)$ eines Bildpunktes und dem entsprechenden Grauwert $f(x,y)$ dar:

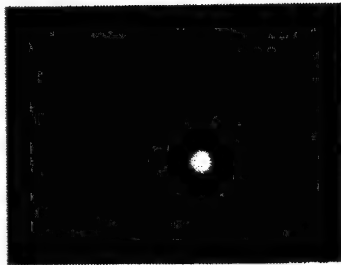
$$f(x,y) = v(x,y) \cdot fGain \cdot 256 + 4,25 \cdot iLevel \quad . \quad (6.1)$$



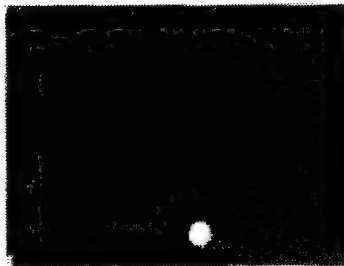
a) $p=500\text{mm}$



b) $p=450\text{mm}$



c) $p=400\text{mm}$



d) $p=350\text{mm}$

Bild 6.2: Bilder mit Laserfleck.

$f\text{Gain}$ ist der Verstärkungsfaktor der Anologschaltung, die das Videosignal vor der Digitalisierung vorverarbeitet; $i\text{Level}$ kontrolliert die Abweichung der Anologschaltung. Wenn man die Werte von $f\text{Gain}$ und $i\text{Level}$ richtig wählt, kann die Fläche der Laserstrahlung im Bild von der übrigen Fläche total getrennt werden.

Durch Experimente sind $f\text{Gain}$ und $i\text{Level}$ als 1.995 (also 4 dB) und -150 ermittelt worden. Bild 6.2b,c, und d zeigen die durch Gleichung (6.1) vorverarbeiteten Bilder. In solchen Bildern sind die Grauwerte des Laserflecks 255. Die übrige Fläche im Bild ist ganz schwarz mit Grauwert 0.

In der Praxis ist das Bild des Laserflecks nicht ein Punktbild, sondern eine unregelmäßige Form. Das macht die Entscheidung so schwierig, welcher Bildpunkt die Koordinaten des Laserflecks im Bild repräsentiert. Zur Vereinfachung wird

der Massenmittelpunkt des Laserflecks als Koordinaten des Laserflecks benutzt, also

$$\begin{cases} x = \frac{\Sigma_x}{F} \\ y = \frac{\Sigma_y}{F} \end{cases} \quad (6.2)$$

F ist die Zahl der Bildpunkte in der Fläche vom Laserfleck im Bild. Σ_x und Σ_y lassen sich durch folgende Gleichung

$$\begin{cases} \Sigma_x = \sum_{x=0}^{511} \sum_{y=0}^{511} f_x(x,y) \\ \Sigma_y = \sum_{x=0}^{511} \sum_{y=0}^{511} f_y(x,y) \end{cases} \quad (6.3)$$

berechnen, mit

$$f_x(x,y) = \begin{cases} x & \text{wenn } f(x,y) = 255 \\ 0 & \text{sonst} \end{cases} ,$$

und

$$f_y(x,y) = \begin{cases} y & \text{wenn } f(x,y) = 255 \\ 0 & \text{sonst} \end{cases} .$$

Die Fläche des Laserflecks F ändert sich, wenn die Kameraposition p' verändert wird. Wenn das Bild genau im Fokus ist, erreicht F das Minimum. Wenn die Kamera von der besten Fokussierungsposition entfernt ist, nimmt F ständig zu. Wie Bild 6.3 zeigt, ist F eine unimodulare Funktion der Kameraposition p' .

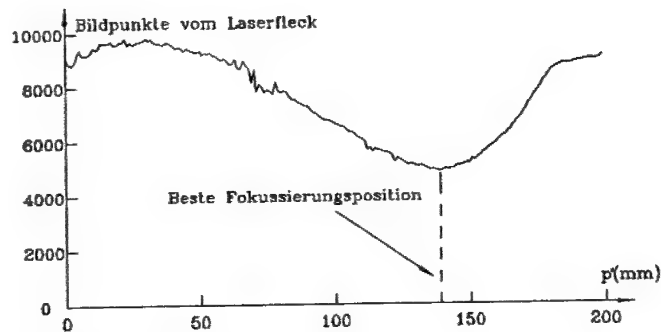


Bild 6.3: Veränderung der Fläche F.

6.2 Beziehung zwischen der Objektweite und den Bildkoordinaten des Laserflecks

Um die Beziehung zwischen der Objektweite p und den Koordinaten des Laserflecks herzuleiten, werden zuerst die Koordinatensysteme festgelegt. Zur Beschreibung der Lage der Objekte gibt es zwei Möglichkeiten (Bild 6.4). Zum einen kann sie in einem Koordinatensystem angegeben werden, das auf die betrachtete Szene bezogen ist, das sogenannte Weltkoordinatensystem (X', Y', Z').

Ein zweites Koordinatensystem ist mit der sich bewegenden Kamera (Kamerakordinatensystem (X, Y, Z)) verknüpft. Physikern sind solche Bewegungsbetrachtungen in verschiedenen Koordinatensystemen, z.B. als ruhender und mitbewegter Beobachter, aus der elementaren Mechanik her vertraut. In der Bildverarbeitung

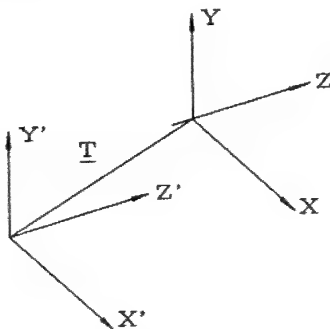


Bild 6.4: Welt- und Kamerakordinatensysteme.

steht anstelle des Beobachters eine Kamera. Es ist allgemeine Konvention, die X- und Y-Achse in die Bildebene der Kamera zu legen, während die Z-Achse mit der optischen Achse des Kamerasystems identisch ist.

Normalerweise sind die beiden Koordinatensysteme nicht koinzidiert. Durch Rotation um den Ursprung und eine Translation mit dem Translationsvektor \underline{I} kann das Weltkoordinatensystem in das Kamerakoordinatensystem übergeführt werden.

In der vorgelegten Arbeit ist das Weltkoordinatensystem fest. Während des Fokussierungsprozesses bewegt sich das Kamerasystem. Deswegen hat das Kamerakoordinatensystem eine relative Bewegung zum Weltkoordinatensystem. Normalerweise weiß man zuerst die Weltkoordinaten des Objektes. Der Übergang von den Welt- zu den Kamerakoordinaten wird durch eine Koordinatentransformation beschrieben. Diese läßt sich in zwei elementare Transformationen aufteilen: Zuerst wird das Kamerakoordinatensystem so gedreht, daß es genauso ausgerichtet ist, wie das Weltkoordinatensystem. Daran anschließend folgt eine Translation des Koordinatensystems um einen Translationsvektor \underline{I} . Die Translationstransformation wird durch folgende Formel zusammengefaßt:

$$\begin{cases} X = X' + T_X \\ Y = Y' + T_Y \\ Z = Z' + T_Z \end{cases} \quad (6.4)$$

Gleichung (6.4) kann als folgende Matrixform umgeschrieben werden:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} + \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \quad (6.5)$$

Die Hintereinanderschaltung mehrerer Transformationen ist manchmal nützlicher als eine einzige Kombinationstransformation. Die mathematische Darstellung dieses Prozesses kann einfach durch eine quadratische Matrix realisiert werden. So kann (6.5) in folgende Gleichung umgeschrieben werden:

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & T_X \\ 0 & 1 & 0 & T_Y \\ 0 & 0 & 1 & T_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} \quad (6.6)$$

Durch die o.g. Umschreibung können alle Transformationen als eine vereinigte Matrix repräsentiert werden

$$\underline{Y} = \underline{A} \cdot \underline{Y'} \quad (6.7)$$

wobei \underline{A} eine 4x4 Transformationsmatrix ist und

$$\underline{Y} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad \underline{Y'} = \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix}$$

Für die Translation ist die Transformationsmatrix durch folgende Form dargestellt:

$$\underline{A}_T = \begin{bmatrix} 1 & 0 & 0 & T_X \\ 0 & 1 & 0 & T_Y \\ 0 & 0 & 1 & T_Z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Die Rotation im dreidimensionalen Raum wird durch eine orthogonale Matrix erzeugt. Eine allgemeine Transformation in ein Koordinatensystem beliebiger Orientierung kann durch drei hintereinander ausgeführte Drehungen um verschiedene Koordinatenachsen zusammengesetzt werden. Die Gesamtdrehung kann man sich als Teildrehung um die X-Achse ($= \tilde{\alpha}$), um die Y-Achse ($= \tilde{\beta}$) und um die Z-Achse ($= \tilde{\theta}$) zusammengesetzt denken.

Die Transformationsmatrix für die Drehung um die X-Achse $R_{\tilde{\alpha}}$ ist

$$R_{\tilde{\alpha}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \tilde{\alpha} & \sin \tilde{\alpha} & 0 \\ 0 & -\sin \tilde{\alpha} & \cos \tilde{\alpha} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Der Rotationswinkel $\tilde{\alpha}$ ist im Uhrzeigersinn gemessen. Diese Transformation beeinflusst nur die Werte der X- und Y-Koordinaten.

Die Teildrehung um die Y-Achse und um die Z-Achse lassen sich durch die Transformationsmatrizen $R_{\tilde{\beta}}$ und $R_{\tilde{\theta}}$ beschreiben:

$$R_{\tilde{\beta}} = \begin{bmatrix} \cos \tilde{\beta} & 0 & \sin \tilde{\beta} & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \tilde{\beta} & 0 & \cos \tilde{\beta} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad R_{\tilde{\theta}} = \begin{bmatrix} \cos \tilde{\theta} & \sin \tilde{\theta} & 0 & 0 \\ -\sin \tilde{\theta} & \cos \tilde{\theta} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Die gesamte Transformation kann in die vier Teiltransformationen A_T , $R_{\tilde{\alpha}}$, $R_{\tilde{\beta}}$ und $R_{\tilde{\theta}}$ zerlegt werden. Die gesamte Transformationsmatrix ist durch das Produkt der Teiltransformationsmatrizen dargestellt.

Durch die Transformationen kann man die Kamerakoordinaten (X,Y,Z) eines Punktes in die Szene bekommen. Die Beziehung zwischen den Kamerakoordinaten (X,Y,Z) und den Bildkoordinaten (x,y,z) ist durch eine perspektivische Transformation beschrieben. Eine perspektivische Transformation projiziert die dreidimensionalen Punkte in eine Bildebene. Die perspektivische Transformation spielt eine zentrale Rolle bei der Bildverarbeitung, weil sie eine Approximation zur Abbildung einer dreidimensionalen Welt in ein Bild ermöglicht.

Als einfachstes Kameramodell wird eine Lochkamera gewählt, deren Abbildungsgeometrie in Bild 6.5 dargestellt ist.

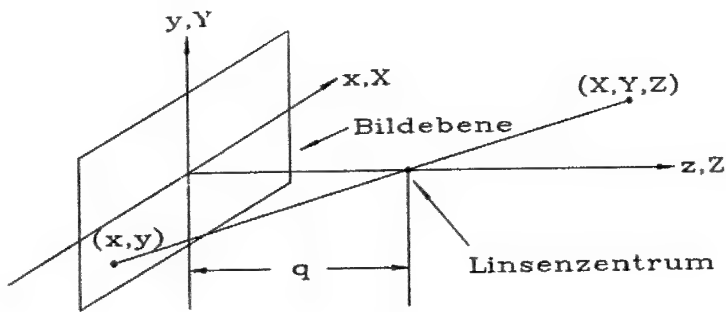


Bild 6.5: Zentralprojektion bei der Abbildung mit einer Lochkamera.

Wie Bild 6.5 zeigt, ist das Zentrum des Bildes der Ursprung des Bildkoordinatensystems und das Linsenzentrum hat die Koordinaten $(0,0,q)$. Das Kamerakoordinatensystem (X,Y,Z) und Bildkoordinatensystem (x,y,z) sind übereinanderliegend.

Gemäß der geometrischen Optik gehen alle Verbindungslinien (Lichtstrahlen) zwischen der Objektebene, die mit den Kamerakoordinaten (X,Y,Z) beschrieben wird, und der Bildebene mit den Bildkoordinaten (x,y,z) durch das Linsenzentrum. Daher können wir den Strahlensatz der Zentralprojektion anwenden, um eine Beziehung zwischen den Kamera- und den Bildkoordinaten zu erhalten.

$$\begin{cases} \frac{x}{q} = -\frac{X}{Z-q} = \frac{X}{q-Z} \\ \frac{y}{q} = -\frac{Y}{Z-q} = \frac{Y}{q-Z} \end{cases} \quad (6.8)$$

Die negativen Vorzeichen vor X und Y bedeuten, daß die Bildpunkte tatsächlich umgekehrt sind. Die Bildkoordinaten des projizierten 3D-Punktes folgen direkt aus (6.8):

$$\begin{cases} x = \frac{q \cdot X}{q-Z} \\ y = \frac{q \cdot Y}{q-Z} \end{cases} \quad (6.9)$$

Es ist sehr wichtig zu bemerken, daß diese Gleichung nichtlinear ist, weil sie eine Division durch die Variable Z enthält. Mit Hilfe einer einzigen Matrix können Translation, Rotation und Projektion als eine verallgemeinerte Transformation beschreiben werden. Zu diesem Zweck werden die homogenen Koordinaten benutzt.

Die homogenen Koordinaten eines Punktes mit den kartesischen Koordinaten (X,Y,Z) sind definiert als (k·X,k·Y,k·Z,k), wobei k eine willkürliche, von Null verschiedene Konstante ist. Die Umwandlung der homogenen Koordinaten zurück zu kartesischen Koordinaten erhält man durch die Teilung der ersten drei homogenen Koordinaten durch die vierte. Ein Punkt im kartesischen Koordinatensystem \underline{v} kann durch den Vektor \underline{v}_h im homogenen System dargestellt werden:

$$\underline{v}_h = \begin{bmatrix} k \cdot X \\ k \cdot Y \\ k \cdot Z \\ k \end{bmatrix} \quad (6.10)$$

Die Transformationsmatrix der perspektivischen Transformation läßt sich definieren zu:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/q & 0 \end{bmatrix} \quad (6.11)$$

Dann liefert das Produkt $P \cdot \underline{v}_h$ einen neuen Vektor, nämlich

$$\underline{c}_h = P \cdot \underline{v}_h = \begin{bmatrix} k \cdot X \\ k \cdot Y \\ k \cdot Z \\ -k \cdot Z/q + k \end{bmatrix} \quad (6.12)$$

Die Komponenten von \underline{c}_h sind die Bildkoordinaten in homogener Form. Die kartesischen Koordinaten von irgendeinem Punkt in dem Kamerakoordinatensystem sind in Vektorform

$$\underline{c} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} q \cdot X / (q - Z) \\ q \cdot Y / (q - Z) \\ q \cdot Z / (q - Z) \end{bmatrix} \quad (6.13)$$

Die ersten beiden Komponenten von \underline{C} sind die (x,y) -Koordinaten in der Bildebene von einem 3D-Punkt (X,Y,Z) . Die dritte Komponente z ist nicht interessant für die Bildverarbeitung, wenn man das Kameramodell aus Bild 6.4 benutzt. Diese Komponente funktioniert als eine freie Variable in der invers perspektivischen Transformation.

Die invers perspektivische Transformation bildet einen Bildpunkt zur 3D-Objektebene zurück. Aus Gleichung (6.12) ergibt sich

$$\underline{v}_h = \underline{P}^{-1} \cdot \underline{C}_h \quad (6.14)$$

Normalerweise sind die Koordinaten vom Bildpunkt durch $(x,y,0)$ dargestellt; $z = 0$ bedeutet, daß die Bildebene sich in der Position $z = 0$ befindet. Dieser Punkt kann in homogener Form in \underline{C}_h durch

$$\underline{C}_h = \begin{bmatrix} k \cdot x \\ k \cdot y \\ 0 \\ k \end{bmatrix} \quad (6.15)$$

repräsentiert werden.

Setzt man \underline{C}_h in Gleichung (6.14) ein, so folgt

$$\underline{v}_h = \begin{bmatrix} k \cdot x \\ k \cdot y \\ 0 \\ k \end{bmatrix}, \quad (6.16)$$

oder, in kartesischen Koordinaten,

$$\underline{v} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \quad (6.17)$$

Dieses Ergebnis ist nicht erwartet, weil es $Z = 0$ für alle 3D-Punkte liefert. Der Grund hierfür ist der folgende: Die Abbildung einer 3D-Szene zu der Bildebene ist eine "viel-zu-eins" Transformation. Der Bildpunkt (x,y) entspricht einer Gruppe der kollinear dreidimensionalen Punkte, die sich auf einer Gera-

den befinden. Diese Gerade geht durch $(x,y,0)$ und $(0,0,q)$. Die Gleichung dieser Geraden im Kamerakoordinatensystem ist:

$$\begin{cases} X = \frac{x}{q} (q-Z) \\ Y = \frac{y}{q} (q-Z) \end{cases} \quad (6.18)$$

Diese Gleichung zeigt: Nur wenn man etwas über die 3D-Punkte weiß, z.B. die Z-Koordinaten, kann man diesen 3D-Punkt aus seinen Bildkoordinaten vollständig zurückgewinnen. Um dieses Problem zu vermeiden, wird die z Komponente von \underline{C}_h nicht als Null, sondern als eine freie Variable behandelt, also

$$\underline{C}_h = \begin{bmatrix} k \cdot x \\ k \cdot y \\ k \cdot z \\ k \end{bmatrix} \quad (6.19)$$

Aus Gleichung (6.14) folgt dann

$$\underline{v}_h = \begin{bmatrix} k \cdot x \\ k \cdot y \\ k \cdot z \\ k \cdot z / (q+k) \end{bmatrix} \quad (6.20)$$

Die kartesischen Koordinaten von \underline{v} sind

$$\begin{cases} X = q \cdot x / (q+z) \\ Y = q \cdot y / (q+z) \\ Z = q \cdot z / (q+z) \end{cases} \quad (6.21)$$

Aus der dritten Gleichung von (6.21) wird z herausgelöst. Man erhält

$$z = \frac{q \cdot Z}{q - Z}$$

Setzt man z in die ersten zwei Gleichungen von (6.21) ein, ergibt sich sofort die Gleichung (6.18). Diese Gleichung bedeutet, daß die Wiederherstellung eines dreidimensionalen Punktes aus seinen Bildkoordinaten durch die invers perspektivische Transformation die Kenntniss über wenigstens eine der drei Kame-

rakoordinaten erfordert. Im beschriebenen System ist das Kamerakoordinatensystem vor dem Fokussierungsprozeß mit dem Weltkoordinatensystem koinzidiert. Während des Fokussierungsprozesses wird sich das Kamerakoordinatensystem nur längs der Z' -Achse des Weltkoordinatensystems bewegen. Es gibt keine Rotation zwischen beiden Koordinatensystemen. Deshalb ist die gesamte Translationsmatrix A_T durch

$$A_T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

dargestellt. Setzt man A_T in (6.7) ein, so folgt

$$\begin{cases} X = X' \\ Y = Y' \\ Z = Z' - T_z \end{cases} \quad (6.22)$$

Nach der obigen Vorbereitung kann man jetzt die Beziehung zwischen den Bildkoordinaten des Laserflecks und der Objektweite herleiten. Zur Vereinfachung wird das Lasergerät so eingestellt, daß die Laserstrahlung parallel zur $X'Z'$ -Ebene des Weltkoordinatensystems ist. Deshalb verändert die Verschiebung des Objektes nur die Z' - und X' -Koordinaten der Laserstrahlung, Y' bleibt unveränderlich.

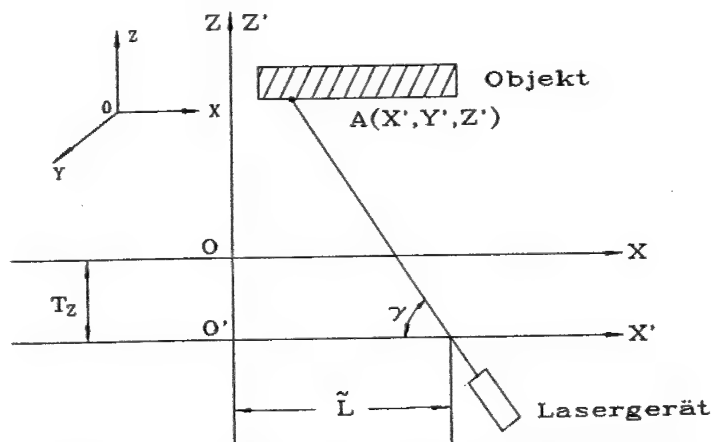


Bild 6.6: Geometrie der dreidimensionalen Koordinaten vom Laserfleck.

Wie Bild 6.6. zeigt, sind der Winkel γ und der Abstand zwischen dem Ursprung O' und dem Lasergerät \tilde{L} nach der Einstellung des Lasergerätes bekannt. Deshalb kann man die Koordinate Z' des Objektes nur bekommen, wenn die Koordinate X' bekannt ist. Aus der Trigonometrie ergibt sich

$$Z' = (\tilde{L} - X') \cdot \operatorname{tg} \gamma \quad (6.23)$$

Aus Gleichung (6.22) erhält man die Kamerakoordinaten der Laserstrahlung

$$\begin{cases} X = X' \\ Y = Y' \\ Z = Z' - Tz = (\tilde{L} - X') \cdot \operatorname{tg} \gamma - Tz = (\tilde{L} - X) \cdot \operatorname{tg} \gamma - Tz \end{cases} \quad (6.24)$$

Während des Fokussierungsprozesses kann die Koordinate X' vom Laserfleck A nicht direkt gemessen werden, man kann aber die Bildkoordinaten (x, y) vom Punkt A bekommen. Setzt man (6.24) in Gleichung (6.18) ein, so folgt

$$x = \frac{x \cdot [q - (\tilde{L} - X) \cdot \operatorname{tg} \gamma + Tz]}{q} \quad (6.25)$$

Aus (6.25) folgt

$$x = \frac{x \cdot (q - \tilde{L} \cdot \operatorname{tg} \gamma + Tz)}{q - x \cdot \operatorname{tg} \gamma} \quad (6.26)$$

Setzt man (6.26) wieder in (6.24) ein, dann ist Z sofort gefunden. Die Koordinate lautet

$$Z = \frac{(\tilde{L} - x) \cdot \operatorname{tg} \gamma - Tz}{q - x \cdot \operatorname{tg} \gamma} \cdot q \quad (6.27)$$

Aus (6.27) wird ersichtlich, wenn $(q - x \cdot \operatorname{tg} \gamma)$ nicht identisch Null ist, kann über die Bildkoordinate x die Kamerakoordinate Z des Laserflecks eindeutig bestimmen werden. $(q - x \cdot \operatorname{tg} \gamma) = 0$ bedeutet, daß die Laserstrahlung durch das Linsenzentrum läuft. Das ist aber unmöglich. Wenn die Kamerakoordinate Z bekannt ist, kann die Kamera direkt zur besten Fokussierungsposition fahren.

Die Besonderheit dieses Fokussierungsalgorithmus liegt darin, daß der Abstand zwischen der Linse und dem Objekt mit Hilfe der Laserstrahlung direkt berechnet werden kann. Die Grauwerte des Bildes spielt dabei keine Rolle. Der Prozeß der Fokussierung eines beweglichen Objektes kann in folgende Schritte eingeteilt werden:

1. Schritt: Der Computer wird die Kamerakoordinate Z des zu fokussierenden Objektes mit Hilfe der Bildkoordinate x des Laserflecks berechnen. Die Objektweite p ist dann $Z \cdot q$.
2. Schritt: Gemäß der Objektweite p beurteilt der Computer, ob die Kamera sich in der besten Fokussierungsposition befindet.
3. Schritt: Die Kamera fährt zur besten Fokussierungsposition.

Während des Fokussierungsprozesses werden die obigen drei Schritte wiederholt, damit das bewegliche Objekt ununterbrochen fokussiert werden kann.

6.3 Kamerakalibrierung und Fehleranalyse

Die Kamerakalibrierung ist ein wichtiger Vorgang für die Anwendungen der Bildverarbeitung. Tsai [28] hat einen umfassenden Überblick über die Kamerakalibrierung gegeben. Die vorhandene Technik der Kamerakalibrierung ist in zwei Kategorien eingeteilt: das Lochkameramodell (lineares Lochkameramodell [29], [30] und nichtlineares Lochkameramodell [31], [32]) und das Zweiebenenmodell [33]. In der vorliegenden Arbeit wird das lineare Lochkameramodell benutzt.

Die Gleichungen (6.14) und (6.15) charakterisieren die Abbildung eines Bildes durch die Projektion der 3D-Punkte in die Bildebene. Diese beiden Gleichungen bilden die Basis für das mathematische Modell einer Kamera. Dieses Modell basiert auf der Annahme, daß die Kamera- und Weltkoordinatensysteme übereinanderliegen. In vielen Situationen sind beide Koordinatensysteme allerdings getrennt. Aber durch eine Rotation um den Ursprung und eine Translation mit dem Translationsvektor \mathbf{I} kann das Kamerakoordinatensystem in das Weltkoordinatensystem übergeführt werden. Um die gesamte Translationsmatrix zu bekommen, muß man die Kenntnisse über die Bildweite q , den Translationsvektor \mathbf{I} und die drei Drehwinkel $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\theta}$ genau kennen.

Diese Größen kann man natürlich direkt messen. Aber es ist häufig günstiger, ihre Bestimmung mit Hilfe der Kamera durchzuführen. In diesem Prozeß wird die Kamera als ein Meßgerät genutzt. Dazu wird eine Gruppe von Bildpunkten und deren Weltkoordinaten benötigt. Der Berechnungsprozeß zum Erhalt der Elemente der Transformationsmatrix mit Hilfe der bekannten Punkte wird häufig als Kamerakalibrierung bezeichnet.

Normalerweise sind die homogenen Bildkoordinaten c_h und die homogenen Weltkoordinaten $(k \cdot X', k \cdot Y', k \cdot Z', k)$ mit $k = 1$ durch folgende Gleichung verbunden

$$\begin{bmatrix} c_{h1} \\ c_{h2} \\ c_{h3} \\ c_{h4} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \cdot \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} \quad (6.28)$$

Wie oben erwähnt, sind die Bildkoordinaten in kartesischer Form durch

$$\begin{cases} x = \frac{c_{h1}}{c_{h4}} \\ y = \frac{c_{h2}}{c_{h4}} \end{cases} \quad (6.29)$$

gegeben. Das Ersetzen von $c_{h1} = x \cdot c_{h4}$ und $c_{h2} = y \cdot c_{h4}$ in Gleichung (6.28) und die Durchführung der Matrizenoperation liefert

$$\begin{cases} x \cdot c_{h4} = a_{11} \cdot X' + a_{12} \cdot Y' + a_{13} \cdot Z' + a_{14} \\ y \cdot c_{h4} = a_{21} \cdot X' + a_{22} \cdot Y' + a_{23} \cdot Z' + a_{24} \\ c_{h4} = a_{41} \cdot X' + a_{42} \cdot Y' + a_{43} \cdot Z' + a_{44} \end{cases} \quad (6.30)$$

In Gleichung (6.30) wird c_{h3} nicht berücksichtigt, weil c_{h3} in Beziehung mit der freien Variable z steht. Das Ersetzen von c_{h4} in den ersten zwei Gleichungen von (6.30) ergibt zwei Gleichungen mit 12 unbekannten Koeffizienten,

$$\begin{cases} a_{11} \cdot X' + a_{12} \cdot Y' + a_{13} \cdot Z' - a_{41} \cdot x \cdot X' - a_{42} \cdot x \cdot Y' - a_{43} \cdot x \cdot Z' - a_{44} \cdot x + a_{14} = 0 \\ a_{21} \cdot X' + a_{22} \cdot Y' + a_{23} \cdot Z' - a_{41} \cdot y \cdot X' - a_{42} \cdot y \cdot Y' - a_{43} \cdot y \cdot Z' - a_{44} \cdot y + a_{24} = 0 \end{cases} \quad (6.31)$$

Der Kalibrierungsprozeß besteht nun aus:

- a) Es werden wenigstens $m = 6$ Punkte mit den Koordinaten (X'_i, Y'_i, Z'_i) , $i = 1, 2, \dots, m$ im Weltkoordinatensystem gemessen.
- b) Die entsprechenden Bildkoordinaten (x_i, y_i) , $i = 1, 2, \dots, m$ werden mit Hilfe der Kamera in einer bestimmten Position bestimmt.
- c) Setzt man diese Koordinaten in Gleichung (6.31) ein, so erhält man $2m$ Gleichungen mit 12 unbekannten Koeffizienten.
- d) Löst man dieses Gleichungssystem auf, erhält man die 12 unbekannten Koeffizienten.

Es gibt viele numerische Algorithmen zur Suche einer optimalen Lösung für ein lineares Gleichungssystem. Ein Beispiel hierfür ist die Arbeit von Noble [34].

Wenn die Welt- und Kamerakoordinatensysteme übereinanderliegen, vereinfacht sich die Gleichung (6.31) zu

$$\begin{cases} X' + \frac{1}{q} \cdot x \cdot Z' - x = 0 \\ Y' + \frac{1}{q} \cdot y \cdot Z' - y = 0 \end{cases} \quad (6.32)$$

Das Gleichungssystem (6.32) besteht aus zwei Gleichungen mit nur einem unbekannten Koeffizienten q . Deshalb braucht man nur eine Weltkoordinate (X', Y', Z') und die entsprechende Bildkoordinate (x, y) . Aus (6.32) erhält man q_1 und q_2

$$\begin{cases} q_1 = \frac{x \cdot Z'}{x - X'} \\ q_2 = \frac{y \cdot Z'}{y - Y'} \end{cases}$$

Bei Anwendung des "Least-Squared Error" Verfahrens kann ein optimales q erzielt werden. Es lautet

$$q = \frac{q_1 + q_2}{2} = \frac{2x \cdot y - x \cdot Y' - y \cdot X'}{2(x - X') \cdot (y - Y')} \cdot Z' \quad (6.33)$$

Außer der Bildweite q sind noch zwei andere Koeffizienten γ und $\tilde{\gamma}$ im System festzulegen. Von (6.27) kann man sehen, um γ und $\tilde{\gamma}$ zu bestimmen, müssen die Kamerakoordinaten von zwei Punkten (X_1, Y_1, Z_1) , (X_2, Y_2, Z_2) und die entsprechenden Bildkoordinaten (x_1, y_1) , (x_2, y_2) bekannt sein. Setzt man diese Koordinaten in (6.27) ein, so ergibt sich

$$\begin{cases} Z_1 \cdot (q - x_1 \cdot \gamma) = q \cdot (\tilde{\gamma} - x_1) \cdot \gamma - q \cdot T_z \\ Z_2 \cdot (q - x_2 \cdot \gamma) = q \cdot (\tilde{\gamma} - x_2) \cdot \gamma - q \cdot T_z \end{cases} \quad (6.34)$$

Löst man dieses Gleichungssystem auf, erhält man γ und $\tilde{\gamma}$ zu

$$\begin{cases} \gamma = \text{tg}^{-1} \frac{q \cdot (Z_1 - Z_2)}{q \cdot (x_2 - x_1) - (x_2 \cdot Z_2 - x_1 \cdot Z_1)} \\ \tilde{\gamma} = \frac{q \cdot x_2 \cdot Z_1 + q \cdot x_2 \cdot T_z + x_1 \cdot Z_1 \cdot T_z - q \cdot x_1 \cdot Z_2 + x_1 \cdot Z_1 \cdot Z_2 - x_2 \cdot Z_1 \cdot Z_2 - q \cdot x_1 \cdot T_z - x_2 \cdot Z_1 \cdot T_z}{q \cdot (Z_1 - Z_2)} \end{cases} \quad (6.35)$$

Wenn die Welt- und Kamerakoordinatensysteme übereinanderliegen, ist $T_z = 0$. Dann vereinfacht sich (6.35) zu

$$\begin{cases} \gamma = \text{tg}^{-1} \frac{q \cdot (Z_1 - Z_2)}{q \cdot (x_2 - x_1) - (x_2 \cdot Z_2 - x_1 \cdot Z_1)} \\ \tilde{\gamma} = \frac{q \cdot x_2 \cdot Z_1 - q \cdot x_1 \cdot Z_2 + x_1 \cdot Z_1 \cdot Z_2 - x_2 \cdot Z_1 \cdot Z_2}{q \cdot (Z_1 - Z_2)} \end{cases} \quad (6.36)$$

Um die drei Koeffizienten q , γ und $\tilde{\gamma}$ zu bekommen, sind zwei Kamerakoordinaten und die entsprechenden Bildkoordinaten erforderlich, also beispielsweise

$$\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} = \begin{bmatrix} -12,89 \\ 1,34 \\ 590 \end{bmatrix}, \quad \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 8,85 \\ 1,35 \\ 455 \end{bmatrix}, \quad \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1,677 \\ -0,174 \end{bmatrix}, \quad \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1,415 \\ -0,224 \end{bmatrix}.$$

Setzt man diese Koordinaten in (6.33) und (6.36) ein, so folgt

$$q = 67,86 \text{ mm} ; \gamma = 81^{\circ}52' ; \ell = 71,53 \text{ mm}.$$

Die durch (6.27) gewonnene Z-Koordinate ist abhängig von den Koeffizienten q , γ , ℓ , Tz und x . Jeder Fehler in diesen Größen wird Fehler von Z verursachen. Die Größen q , γ , ℓ werden vor dem Fokussierungsprozeß bestimmt und während des Prozesses nicht verändert. Die x -Koordinate wird durch viele Faktoren beeinflusst und während des Prozesses laufend gemessen. Der Fehler von x ist sehr schwer zu kontrollieren. Tabelle 6.1 zeigt die realen und die durch (6.27) berechneten Werte von Z .

Nr.	$Z_R(\text{mm})$	$Z_M(\text{mm})$	Relativer Fehler (%)
1	380	381,7	0,45
2	390	391,8	0,45
3	400	398,2	0,45
4	410	408,4	0,40
5	420	418,3	0,41
6	430	427,4	0,60
7	440	437,3	0,61
8	450	452,7	0,60
9	460	462,4	0,52
10	470	471,8	0,40
11	480	482,5	0,52
12	490	492,2	0,45
13	500	502,4	0,48

Tab. 6.1: Fehler von Z .

Gemäß der Fehlerfortpflanzung kann der absolute Fehler δZ von Z aus Gleichung (6.27) hergeleitet werden:

$$(\delta Z)^2 = \left(\frac{\partial Z}{\partial q} \cdot \delta q\right)^2 + \left(\frac{\partial Z}{\partial \ell} \cdot \delta \ell\right)^2 + \left(\frac{\partial Z}{\partial \gamma} \cdot \delta \gamma\right)^2 + \left(\frac{\partial Z}{\partial Tz} \cdot \delta Tz\right)^2 + \left(\frac{\partial Z}{\partial x} \cdot \delta x\right)^2 . \quad (6.37)$$

Aus (6.27) folgt dabei:

$$\begin{aligned}\frac{\partial Z}{\partial q} &= \frac{-x \cdot \operatorname{tg} \gamma \cdot [(\tilde{l} - x) \cdot \operatorname{tg} \gamma - T_z]}{(q - x \cdot \operatorname{tg} \gamma)^2}, \\ \frac{\partial Z}{\partial \tilde{l}} &= \frac{\operatorname{tg} \gamma \cdot q}{q - x \cdot \operatorname{tg} \gamma}, \\ \frac{\partial Z}{\partial \gamma} &= \frac{(q - x \cdot \operatorname{tg} \gamma) \cdot \sec^2 \gamma + x \cdot q \cdot [(\tilde{l} - x) \cdot \operatorname{tg} \gamma - T_z] \cdot \sec^2 \gamma}{(q - x \cdot \operatorname{tg} \gamma)^2}, \\ \frac{\partial Z}{\partial T_z} &= \frac{-q}{q - x \cdot \operatorname{tg} \gamma}, \\ \frac{\partial Z}{\partial x} &= \frac{q \cdot (\tilde{l} \cdot \operatorname{tg} \gamma - q - T_z) \cdot \operatorname{tg} \gamma}{(q - x \cdot \operatorname{tg} \gamma)^2}.\end{aligned}$$

Die Bewegung der Kamera verändert den Wert von T_z . Weil die Bewegung der Kamera durch einen Schrittmotor realisiert ist, kann die Position der Kamera genau bestimmt werden, also kann δT_z vernachlässigt werden. Die Fehler $\delta \gamma$, $\delta \tilde{l}$ und δq kann man im Prozeß der Kamerakalibrierung kontrollieren. Der Fehler δx ist deshalb die Hauptursache für den Fehler δZ .

7. Entwurf eines Fokussierungsregelkreises

7.1 Max Vision AT-1 Bildverarbeitungssystem

Für die Entwicklung der automatischen Fokussierung der Kamera wurde das Bildverarbeitungssystem Max Vision AT-1 der Firma DATACUBE eingesetzt. Im Max Vision System steht eine Reihe von Geräten zur Verfügung. Bild 7.1 ist eine schematische Darstellung des Systems.

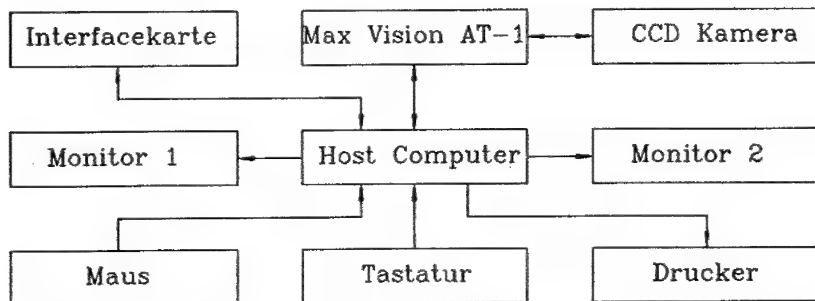


Bild 7.1: Blockschaltbild der Hardware-Konfiguration des Bildverarbeitungssystems.

1) Der Host Computer

Der Host Computer (ESCOM 386/33) ist ein Personal Computer, der mit dem Computer IBM AT kompatibel ist. Der Host Computer dient zur Durchführung der Bildverarbeitungsprogramme und zur Steuerung aller Peripherien. Das Betriebssystem von ihm ist MS-DOS.

2) Das Max Vision AT-1 System

Das Max Vision System ist ein Gerät, das die komplizierten Aufgaben der Bildverarbeitung einfach lösen kann. Die kräftige Hardware und extensive Software-Bibliothek vom Max Vision System macht es geeignet zur Unterstützung verschiedener Anwendungen.

Das Max Vision System arbeitet in Verbindung mit dem Host Computer und hat einen sogenannten "Pipelined Digital Processor" (PDP) mit sehr schneller Geschwindigkeit. Der PDP ist ein digitaler Matrixprozessor, der aus einem Multiplier, zwei Look-Up-Tabellen für Operanden, einer Arithmetik-Logikeinheit

(ALU) mit zwei 16-Bit Operanden und einer Clipper- und Shifterschaltung für die Datennormalisierung besteht. Zur Ein- und Ausgabe des Bildes werden im Max Vision System ein A/D und ein D/A Converter verwendet. Bis zu 4 Videokameras können mit dem Max Vision System verbunden werden. Außerdem kann das System gleichzeitig 3 Bilder in inneren Bildspeicher speichern.

Der PDP kann folgende Operationen durchführen: Transformation der Daten zwischen zwei Look-Up-Tabellen, Multiplikationsoperationen, Logikoperationen, Arithmetikoperationen, Datennormalisierung für alle 262144 (512x512) Bildpunkte in Echtzeit (innerhalb 1/30 Sekunde).

3) CCD-Kamera

Die Quelle des Videosignals kommt aus einer selbst aufgebauten CCD-Kamera. Diese CCD-Kamera erzeugt das CCIR 50 Hz standardisierte analoge Videosignal. Der Kern der CCD-Kamera ist ein monochromatischer Abbildungsmodul NCA1101/01 der Firma Philips. Die Bildfläche des Abbildungsmoduls ist 6,00 mm (horizontal) X 4,5 mm (vertikal). Die Zahl der Bildpunkte beträgt 604 x 588. Das Max Vision System benutzt davon nur 512 x 512 Bildpunkte. Die Größe jedes Bildpunktes ist deshalb ungefähr 0,01 mm x 0,00773 mm.

4) Die Monitore

Es gibt zwei Monitore im System. Ein Farbmonitor dient als Ausgabeeinheit vom Bild. Er kann gleichzeitig drei (rot, grün, blau) analoge Videosignale übernehmen. Der andere Monitor arbeitet als Kontrollbildschirm für den Host Computer.

5) Die "ISERT" Interfacekarte 4.0

Die "ISERT" Interfacekarte 4.0 ist ein Schrittmotorantrieb, der in den letzten Jahren im Bereich der Positionierungssteuerung und Koordinatenantriebe immer mehr an Bedeutung gewonnen hat. Der Schrittmotor wandelt elektrische Impulse aus der Interfacekarte in eine mechanische Drehbewegung mit definiertem Schrittwinkel um. Jedes Steuersignal aus der Interfacekarte 4.0 erzeugt einen Schritt für den Schrittmotor.

Die Interfacekarte 4.0 ist ein mikroprozessorgestütztes Steuerungssystem für Schrittmotoransteuerungen. Die charakteristische Eigenschaft dieser Antriebe ist in dem direkten Zusammenhang zwischen den elektrischen Signalen aus dem Host Computer und der Drehbewegung des Schrittmotors zu sehen. Durch die aus-

gereifte intelligente Steuerungssoftware in Verbindung mit einem schnellen 11 MHz-Mikrocontroller ist sie einfach zu programmieren.

Die Interfacekarte unterstützt die Ansteuerung von insgesamt drei angeschlossenen Schrittmotoren. Dabei werden aus den übergebenen Informationen wie Bewegungsabstand und Geschwindigkeitssollwert sowie aus den intern festgelegten Antriebsparametern die erforderlichen Werte für die Start- und Stoprampe berechnet.

In dem angeschlossenen Computer werden die Befehle für die Bearbeitung der notwendigen Einzelschritte erfaßt. Die erfaßten Befehle werden über die serielle Verbindung zwischen Computer und Interfacekarte an die Schrittmotorsteuerung übergeben. Diese Befehle werden von der Schrittmotorsteuerung direkt ausgeführt oder gespeichert. Im beschriebenen System werden alle Befehle direkt ausgeführt. Nach dem Empfang der Befehle arbeitet die Interfacekarte autonom.

Der Host Computer dient als eine Kontrolleinheit. Er erzeugt die Befehle nach Anforderung des Systems und gibt diese Befehle zur Interfacekarte weiter.

6) Der Drucker

Der Drucker dient zum Druck des Bildes auf Papier.

Die kräftige, aber auch komplizierte Hardware von Max Vision wird durch eine mächtige und ziemlich einfache Anwendungssoftware unterstützt. Das Hauptprogramm der Software hat eine Dialogform. Nach der Durchführung des Hauptprogramms wird ein Kontrollmenü auf dem Monitor gezeigt. Auf diesem Menü kann man verschiedene Kommandos wählen. Das Max Vision System kann von drei verschiedenen Niveaus aus kontrolliert werden.

1. "Menu Driven Interface" Niveau

Bei diesem Niveau können die Benutzer verschiedene Bildverarbeitungs Routinen aus verschiedenen Kontrollmenüs benutzen. Die Ergebnisse der Verarbeitung werden sofort auf dem Monitor gezeigt. Man braucht keine Kenntnisse über die Programmiersprache.

2. "Script Operation Interface" Niveau

Das zweite Kontrollniveau zeigt einen sehr einfachen Weg zur Erzeugung vieler anderer Bildverarbeitungskommandos. Bei diesem Niveau kann man viele Funktionen aus der Programmbibliothek benutzen.

3. "C Subroutine Library Interface" Niveau

Bei diesem niedrigsten Kontrollniveau werden Kenntnisse über die Programmiersprache C gebraucht. Die Benutzer können viele Anwendungsprogramme mit C direkt schreiben. Alle Unterprogramme zur automatischen Fokussierung sind direkt in C geschrieben und im Hauptprogramm eingefügt.

7.2 Struktur von Fokussierungsregelkreisen

Im Bildverarbeitungssystem erfüllt die automatische Fokussierung der Kamera zwei Aufgaben: die Erreichung und Erhaltung der besten Fokussierung der Kamera. Bei der Aufgabe zur Erreichung der besten Fokussierung sucht die Kamera in einem großen Arbeitsbereich die beste Fokussierungsposition. Bild 7.2 zeigt die Grundidee zur Lösung dieses Problems.

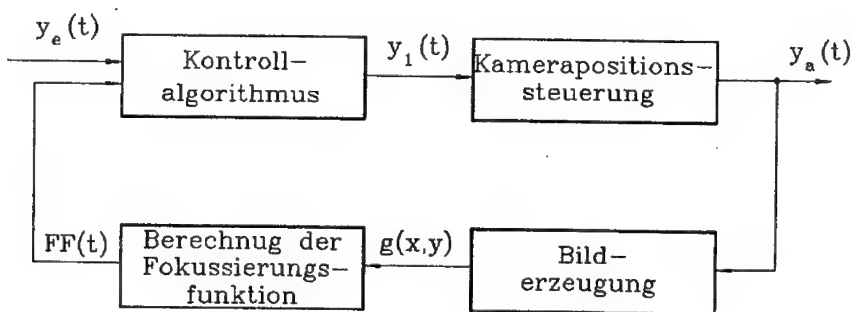


Bild 7.2: Struktur des geschlossenen Regelkreises zur Erreichung der besten Fokussierung.

Für diesen Regelkreis ist der Istwert $y_a(t)$ die Position der Kamera $p'(t)$. Am Ende des Fokussierungsprozesses befindet sich die Kamera in der besten Fokussierungsposition. Deshalb ist der Sollwert $y_e(t)$ die beste Fokussierungsposition der Kamera p'_{opt} . Man weiß, daß das Objekt in einer gegebenen Position nur eine beste Fokussierungsposition der Kamera besitzt. In Bild 3.1 ist p'_{opt}

nur von der Position des Objektes p'_0 abhängig. Vor dem Fokussierungsprozeß ist der Sollwert $y_e(t)$ unbestimmt, weil die Position des Objektes p'_0 unbekannt ist. Der Sollwert von dem Fokussierungsregelkreis ist also nur ein fiktiver Begriff.

Der Block Kamerapositionssteuerung besteht aus der Interfacekarte, dem Schrittmotor und dem Spindelschlittensystem. Gemäß des Signals $y_1(t)$ steuert dieser Block die Bewegung der Kamera. Das Meßglied besteht aus zwei Teilen: Bilderzeugung und Bewertung des Fokussierungszustandes. In der Position y_a bekommt die Kamera ein Bild vom Objekt. Wenn das Objekt unbeweglich ist und die Parameter der Kamera festgelegt sind, ist die Bildqualität nur von y_a abhängig. Anschließend wird der Fokussierungszustand durch die entsprechende Fokussierungsfunktion FF bewertet. Gemäß der Fokussierungsfunktion FF gibt der Regler das Signal $y_1(t)$ aus, das die Bewegung der Kamera kontrolliert.

Zur Analyse der Regelung im Sinne eines geschlossenen Regelkreises werden die Antworten auf folgende Fragen gebraucht:

- a. Wie ist die Beziehung zwischen dem Bild $g(x,y)$ und dem Istwert des Systems y_a . Wenn das Objekt bestimmt und unbeweglich ist, ist die Schärfe des Bildes von y_a abhängig. Das schärfste Bild tritt in der Position $y_a = p'_{opt}$ auf.
- b. Wie ist die Fokussierungsfunktion FF vom Bild $g(x,y)$ und vom Istwert y_a abhängig? Die Fokussierungsfunktion FF ist in den meisten Fällen eine unimodulare Funktion vom Istwert y_a . Der Extremwert der Fokussierungsfunktion tritt in der Position p'_{opt} auf. Die Form von $FF(y_a)$ ist abhängig von der Textur des Objektes. In Wirklichkeit ist $FF(y_a)$ eine hochgradig nichtlineare Funktion von y_a .
- c. Wie ist die Kameraposition (Istwert) y_a von dem Signal y_1 abhängig? Es ist dies eine Funktion der Mechanik und Elektronik vom Block Kamerapositionssteuerung.
- d. Wie ist das Kontrollsignal y_1 von dem Sollwert y_e und der Fokussierungsfunktion $FF(y_a)$ abhängig? Weil y_e nur ein fiktiver Begriff ist, kann der Regler y_e nicht direkt benutzen. Deshalb ist das Kontrollsignal nur von der Fokussierungsfunktion $FF(y_e)$ abhängig.

Um den Fokussierungsregelkreis zu analysieren, ist die sogenannte "Hill-Climbing"-Strategie nützlich [18]. Weil das Kontrollsignal y_{1n} nur die Fokussierungsfunktion $FF(y_a)$ benutzt, ist es unmöglich, die Bewegungsrichtung der Kamera nur mit Hilfe der gegenwärtigen Fokussierungsfunktion festzulegen. Um die Bewegungsrichtung der Kamera zu bekommen, sind die frühere und gegenwärtige Fokussierungsfunktion FF_{n-1} bzw. FF_n notwendig. Ein Kontrollalgorithmus ist also von der Form:

$$y_{1n} = F(FF_n, FF_{n-1}) \quad (7.1)$$

y_{1n} ist das gegenwärtige Kontrollsignal. In der vorliegenden Arbeit ist y_{1n} nur eine Funktion von $FF_n - FF_{n-1}$, also

$$y_{1n} = F(FF_n - FF_{n-1}) \quad (7.2)$$

Nun ist die Fokussierungsstrategie zu der sogenannten "Hill-Climbing"-Strategie entartet. Während des Fokussierungsprozesses wird die Fokussierungsfunktion berechnet, dann bewegt sich die Kamera. In der neuen Position wird die Fokussierungsfunktion erneut gemessen. Anschließend wird die Veränderung der Fokussierungsfunktion geprüft. Wenn die Fokussierungsfunktion abnimmt, bedeutet das, daß die Bewegungsrichtung falsch ist. Die Kamera muß zurückfahren, sonst ist die Bewegungsrichtung der Kamera richtig. In der besten Fokussierungsposition wird die Kamera stehenbleiben.

Jetzt kann man die Funktion F so bestimmen, daß, wenn $FF_n - FF_{n-1} > 0$ ist, y_{1n} positiv ist. Sonst bekommt y_{1n} einen negativen Wert. Ein Regler mit solcher Eigenschaft ist der sogenannte Zweipunktschalter, dessen Kennlinie in Bild 7.3 gezeigt ist.

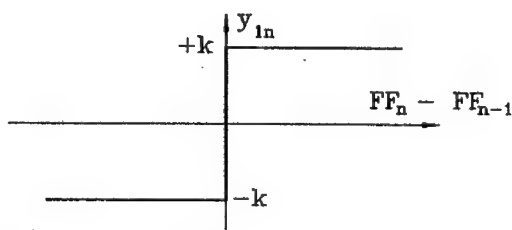


Bild 7.3: Kennlinie des Zweipunktschalters.

Das Kontrollsignal ist dann

$$y_{1n} = k \cdot \text{sgn}(FF_n - FF_{n-1}) \quad (7.3)$$

k ist die Verstärkung. Die "Hill-Climbing"-Fokussierungsstrategie benutzt in der vorliegenden Arbeit eine konstante Schrittgröße des Schrittmotors anstatt der konstanten Verstärkung, d.h. die Kamera fährt nach der Berechnung der Fokussierungsfunktion in einem festen Abstand in die richtige Richtung.

Die Auswahl der Schrittgröße ist sehr wichtig für die Fokussierungsgeschwindigkeit und die Zuverlässigkeit der Fokussierung. Wenn sie zu groß ist, kann die Kamera zwar sehr schnell in die Nähe der besten Fokussierungsposition fahren, die Kamera schwingt aber in der Nähe der besten Fokussierungsposition um diese Position. Mit einer zu kleinen Schrittgröße andererseits findet die Kamera die beste Fokussierungsfunktion zwar zuverlässig, die Fokussierungszeit ist aber zu lang. Um das Problem zu optimieren, wird der Fokussierungsprozeß in drei Phasen aufgeteilt.

a) Erste Phase: grobe Fokussierung.

In dieser Phase braucht die Kamera die beste Fokussierungsposition nur grob zu suchen. Deshalb ist die Schrittgröße sehr groß gewählt. Die Schrittgröße in dieser Phase ist 10 mm. Am Anfang des Fokussierungsprozesses ist die Kamera normalerweise sehr weit vom Objekt entfernt. Die Kamera befindet sich im unwirksamen Bereich. In diesem Bereich ist die Fokussierungsfunktion wegen des Rauscheinflusses nicht immer monoton zunehmend oder abnehmend. Am Ende der ersten Phase wird die Kamera bis zum Punkt, wo die Fokussierungsfunktion die Bedingung

$$FF_n - FF_{n-1} < 0 \quad (7.4)$$

erfüllt ist, fahren. Im idealen Fall, d. h. das Bild ist nicht durch das Rauschen verfälscht, ist die Kamera jetzt die beste Fokussierungsposition überwunden. Aber wegen des Rauscheinflusses ist die FF nicht mehr unimodular. In diesem Fall kann die Bedingung (7.4) im unwirksamen Bereich auch erfüllt werden. Deshalb muß die Bedingung (7.4) verändert werden gemäß

$$FF_n - FF_{n-1} < -\Delta F \quad (7.5)$$

Δf ist eine positive Konstante, die experimentell zu bestimmen ist. Normalerweise ist die Veränderung der Fokussierungsfunktion im unwirksamen Bereich sehr klein im Vergleich zu dem Bereich in der Nähe der besten Fokussierungsposition.

b) Zweite Phase.

In dieser Phase ist die Schrittgröße der Kamera auf 5 mm gesetzt. Die Kamera befindet sich im wirksamen Bereich. Weil am Ende der ersten Phase die Kamera die beste Fokussierungsposition überschreitet, muß die Kamera zurückfahren. Die zweite Phase endet in der Position, wo die Bedingung (7.4) erneut erfüllt ist.

c) Dritte Phase: genaue Fokussierung.

Am Ende der zweiten Phase liegt die Kamera ganz in der Nähe der besten Fokussierungsposition. In dieser Phase fährt die Kamera in der umgekehrten Richtung zur zweiten Phase mit der Schrittgröße 1 mm. Wenn die Kamera die beste Fokussierungsposition nochmals überwindet, fährt die Kamera 1 mm zurück und hat die beste Fokussierungsposition gefunden. Der Fokussierungsprozeß zur Erreichung der besten Fokussierung ist dann beendet.

Der Computer als Regler hat folgende Aufgaben:

1. Berechnung der Fokussierungsfunktion,
2. Festlegung der Fokussierungsphase,
3. Festlegung der Bewegungsrichtung der Kamera,
4. Ermittlung der besten Fokussierungsposition.

7.3 Übertragungsfunktion des Fokussierungsregelkreises

Die Aufgabe des Reglers zur Aufrechterhaltung der besten Fokussierung ist die Suche der besten Fokussierungsposition im kleinen Bereich. Die Voraussetzung dafür ist, daß die Kamera sich in der Nähe der besten Fokussierungsposition befindet. Nach dem Prozeß zur Erreichung der besten Fokussierung liegt die Kamera in der besten Fokussierungsposition. Wenn das Objekt sich in einem kleinen Bereich verschiebt oder die Parameter der Kamera sich verändern, ist die Kamera vom besten Fokussierungszustand abgewichen. Mit Hilfe des Reglers zur Erhaltung der besten Fokussierung kann die Kamera die beste Fokussierungs-

position wiederfinden. Weil die Kamera sich in der Nähe der besten Fokussierungsposition befindet, kann dieser Regelkreis durch die Systemlinearisierung als ein lineares System behandelt werden.

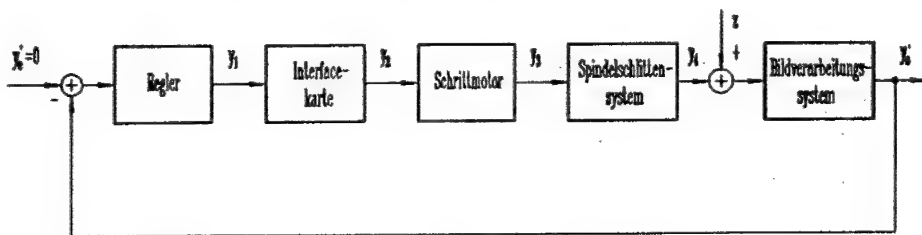


Bild 7.4: Fokussierungsregelkreis zur Erhaltung der besten Fokussierung.

In Bild 7.4 wird der Sollwert des Systems y_e' als das Differential der Fokussierungsfunktion gewählt. Weil die Fokussierungsfunktion den Extremwert in der besten Fokussierungsposition erreicht hat, ist ihr Differential in dieser Position immer Null. Deshalb ist der Sollwert immer 0. Die Übertragungsfunktion der Meßeinrichtung ist 1. Wenn das System stabil ist, ist der Istwert y_a' gleich dem Sollwert y_e' , also die Kamera liegt in der besten Fokussierungsposition.

Die Verschiebung des Objektes und die Veränderung der Parameter kann man als Systemstörung behandeln. Deshalb ist die Hauptfunktion des Regelkreises in Bild 7.4 die Unterdrückung der Systemstörung. Das Störsignal $z(t)$ tritt vor den Block Bildverarbeitungssystem. Es verändert den Abstand zwischen der Kamera und dem Objekt. Die physikalische Bedeutung von $z(t)$ kann so erklärt werden, daß das Objekt plötzlich um einen Abstand $z(t)$ verschoben wird.

1. Übertragungsfunktion der Interfacekarte $G_1(s)$.

Die Interfacekarte im System kann als ein Totzeitglied behandelt werden. Die Totzeit T ist die Zeit zur Behandlung der Anweisung vom Host Computer. Durch das Experiment kann man die Totzeit T messen. Sie beträgt

$$T = 0,3 \text{ (Sekunde)}$$

Unter einer Totzeit versteht man die Erscheinung, daß eine zeitlich veränderliche Größe, die am Eingang eines Blockes wirkt, an dessen Ausgang erst nach einer gewissen Zeit, aber sonst unverändert, in Erscheinung tritt. Mathematisch formuliert heißt das

$$y_2(t) = y_1(t-T) \quad (7.6)$$

Die Übertragungsfunktion eines Totzeitgliedes mit der Totzeit T ergibt sich direkt nach dem Verschiebungssatz der Laplace-Transformation zu

$$G_1(s) = e^{-T \cdot s} \quad (7.7)$$

Es gilt, die transzendente Funktion (7.7) durch ein Polynom oder den Quotienten zweier Polynome anzunähern. Die Funktion (7.7) hat für alle Frequenzen den Betrag 1; das Argument nimmt linear mit der Frequenz zu. Soll die approximierete rationale Funktion ebenfalls für alle Frequenzen den Betrag 1 haben, so müssen Zählerpolynom und Nennerpolynom zueinander konjugiert komplex sein.

Die Funktion ist also von der Form

$$G_1(s) = \frac{a_0 - a_1 \cdot s \cdot T + a_2 \cdot (s \cdot T)^2 - + \dots}{a_0 + a_1 \cdot s \cdot T + a_2 \cdot (s \cdot T)^2 + \dots} \quad (7.8)$$

Die Werte der Koeffizienten a_0, a_1, \dots sind nun so zu wählen, daß der gewünschte lineare Verlauf des Winkels über den interessierenden Frequenzbereich möglichst gut angenähert wird.

Die Entwicklung der Funktion (7.7) in ein Polynom (7.8) unter diesen Gesichtspunkten führt auf die sogenannte Pade-Approximation. Zur Vereinfachung wird das Totzeitglied hier durch die Pade-Approximation erster Ordnung angenähert [36], [37]

$$G_1(s) = \frac{1 - 0,5 \cdot T \cdot s}{1 + 0,5 \cdot T \cdot s} \quad (7.9)$$

2) Übertragungsfunktion des Schrittmotors $G_2(s)$.

Zur Ansteuerung mit Prozeßrechnern oder digitalen Reglern ist der Schrittmotor besonders geeignet. Der Schrittmotor hat ein zu jedem Impuls proportionales Verhalten und kann unter Einsparung eines A/D Wandlers "direkt" digital angesteuert werden [38, 39]. Der angeführte Vorteil für die Positioniersteuerungen bezieht sich auf das einfache Steuerkonzept. Der Antrieb kann lagegesteuert betrieben werden, d.h. er benötigt keine Lageregeleinrichtung und keine Geschwindigkeitsregelung. Die Lage wird durch die Anzahl der Impulse gesteuert und die Geschwindigkeit durch die Schrittfrequenz. Daher kann der Schrittmotor als ein Proportionalglied

$$G_2(s) = k_2 \quad (7.10)$$

betrachtet werden. Der Computer als Regler gibt den Befehl aus. Die Interfacekarte wandelt diesen Befehl in das Geschwindigkeitssignal y_2 um und kontrolliert so die Geschwindigkeit des Schrittmotors. Der Schrittmotor bewegt sich mit der Geschwindigkeit $y_3 = y_2$, deshalb ist die Konstante $k_2 = 1$.

3) Übertragungsfunktion vom Spindelschlittensystem $G_3(s)$.

Das Spindelschlittensystem wandelt die Drehung des Schrittmotors in eine Linearbewegung des Schlittens um. Es ist ein Integrationsglied mit der Übertragungsfunktion

$$G_3(s) = \frac{k_3}{s} \quad (7.11)$$

Wenn sich der Schrittmotor 400 Schritte bewegt, fährt das Spindelschlittensystem 1 mm. Deswegen ist die Konstante k_3

$$k_3 = \frac{1}{400} \text{ (mm/Schritt)} \quad .$$

4) Übertragungsfunktion vom Bildverarbeitungssystem $G_4(s)$.

Der Eingang dieses Übertragungsgliedes ist die Position der Kamera $p'(t)$. Der Ausgang ist das Differential der Fokussierungsfunktion, also

$$y_a'(t) = - \frac{dFF(t)}{dp'(t)} \quad (7.12)$$

Es wird angenommen, daß sich vor dem Auftreten einer Störung die Kamera in der besten Fokussierungsposition befindet. Die beste Fokussierungsposition wird dann als der Ursprung von $p'(t)$ benutzt.

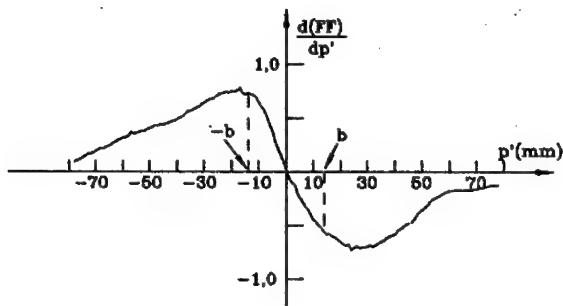


Bild 7.5: Differential der Fokussierungsfunktion.

Die Beziehung zwischen $y'_a(t)$ und $p'(t)$ ist im großen Bereich total nichtlinear (Bild 7.5). Wenn der Absolutwert der Störung kleiner als b ist, also die Kamera in der Nähe vom Ursprung arbeitet, kann die Beziehung zwischen $y'_a(t)$ und $p'(t)$ durch Linearisierung als linear angenommen [40], [41] werden. Dann ist das Bildverarbeitungssystem durch ein lineares Übertragungsglied angenähert. Der Arbeitspunkt für die Linearisierung ist der Ursprung. Im Bereich $(-b, b)$ gilt

$$y'_a(t) = - \frac{dFF(t)}{dp'(t)} = k_4 \cdot p'(t) \quad (7.13)$$

k_4 ist hier eine positive Konstante. Deshalb ist die Übertragungsfunktion $G_4(s)$

$$G_4(s) = k_4 \quad (7.14)$$

Die Bestimmung von k_4 ist sehr kompliziert. Normalerweise ist die Konstante k_4 von Objekt zu Objekt verschieden. Wenn alle Parameter der Kamera festgelegt sind, ist k_4 nur von der Textur des Objektes abhängig. Im Prozeß vor Erreichen der besten Fokussierung kann k_4 gemessen werden.

Weil die Hauptaufgabe vom Fokussierungsregelkreis in Bild 7.4 die Unterdrückung von Störungen ist, soll nun untersucht werden, wie sich die Störung $z(t)$ auf den Istwert $y'_a(t)$ auswirkt. Da ein lineares System vorliegt, läßt sich der Einfluß der Störung $z(t)$ auf $y'_a(t)$ getrennt von dem von $y'_e(t)$ behandeln. In Wirklichkeit ist $y'_e(t)$ eh immer gleich 0. Für die Störung $z(t)$ ergibt sich die Übertragungsfunktion

$$Y'_a(s) = G_z(s) \cdot Z(s) \quad ,$$

$$G_z(s) = \frac{G_4(s)}{1 + G_r(s) \cdot G_1(s) \cdot G_2(s) \cdot G_3(s) \cdot G_4(s)} \quad . \quad (7.15)$$

Setzt man die konkreten Funktionen für $G_1(s)$, $G_2(s)$, $G_3(s)$, $G_4(s)$ in (7.15) ein, so folgt

$$G_z(s) = \frac{400 \cdot k_A \cdot s \cdot (1 + 0,15 \cdot s)}{400 \cdot s \cdot (1 + 0,15 \cdot s) + k_A \cdot G_r(s) \cdot (1 - 0,15 \cdot s)} \quad . \quad (7.16)$$

7.4 Reglerentwurf

Die Aufgabe des Entwurfs vom Regler besteht zunächst darin, einen geeigneten Regler zu wählen. Insbesondere ist zu untersuchen, welcher Regler für bestimmte Regelungsaufgaben vorteilhaft ist und wie man die Parameter des Reglers wählt, damit der Regelkreis die an ihn gestellten Anforderungen erfüllt [42].

Die erste und wichtigste Anforderung an den Regler ist die Stabilität des Regelkreises. Nach Verstellung des Sollwertes $y'_e(t)$ oder der Störung $z(t)$ muß der Istwert mit wachsender Zeit wieder einen festen Wert annehmen, sonst ist der Regelkreis unbrauchbar.

Die zweite Anforderung an den Regler ist, daß der Regelkreis eine bestimmte stationäre Genauigkeit aufweisen muß. Ist der Regelkreis stabil, so strebt der Istwert $y'_a(t)$ nach einer sprunghaftigen Änderung des Sollwertes $y'_e(t)$ oder der Störung $z(t)$ mit wachsender Zeit einem festen Wert zu. Es genügt aber nicht, daß dies irgendein fester Wert ist. Vielmehr muß der Istwert den gewünschten Sollwert annehmen oder ihm zumindest nahekommen.

Die dritte Anforderung wird an die Dynamik des Regelkreises gestellt. Die Antwort des Regelkreises auf einen Führungsgrößensprung muß genügend schnell, aber auch genügend gedämpft sein.

Außerdem muß die Struktur des Reglers so einfach wie möglich sein, damit der Regler leicht realisiert werden kann.

Für den Regelkreis zur Erreichung der besten Fokussierung in Bild 7.2 ist der Regler als Zweipunktschalter ausgewählt worden. Der einzige Parameter ist die Konstante ΔF in Gleichung (7.5). Weil der Regelkreis total nichtlinear ist, wird diese Konstante durch die Experimente bestimmt. Im nächsten Abschnitt wird das Zeitverhalten des Regelkreises für verschiedene Werte ΔF durch Experimente untersucht.

1) PI-Regler

Als Regler wird zuerst ein PI-Regler mit der Übertragungsfunktion

$$G_r(s) = k_r \cdot \frac{1 + T_r \cdot s}{s} \quad (7.17)$$

verwendet. Der PI-Regler besitzt das günstige Zeitverhalten des P-Reglers. Der P-Anteil des Reglers sorgt für eine hinreichende Schnelligkeit. Mit seinem I-Anteil behebt er dessen nachteilige bleibende Regelabweichung.

Bei der in Bild 7.4 gezeichneten Konfiguration lautet die Übertragungsfunktion des offenen Kreises

$$\begin{aligned} G_0(s) &= G_r(s) \cdot G_1(s) \cdot G_2(s) \cdot G_3(s) \cdot G_4(s) \quad , \\ &= k \cdot \frac{(1 - 0,15 \cdot s) \cdot (1 + T_r \cdot s)}{s^2 \cdot (1 + 0,15 \cdot s)} \quad , \end{aligned} \quad (7.18)$$

wobei die Abkürzung

$$k = \frac{k_r \cdot k_4}{400} \quad (7.19)$$

eingeführt wird. Wie ist die Zeitkonstante T_r des Reglers zu wählen und welchen Wert muß die Verstärkung k_r des Reglers haben, damit der Regelkreis befriedigendes Störverhalten aufweist, also insbesondere stabil ist? Dies soll durch Anwendung des Kriteriums von Hurwitz oder des Nyquist-Kriteriums untersucht werden.

- a) In dem ersten Versuch wird der über 0,15 auftretende Pol von (7.18) kompensiert. Setzt man $T_r = 0,15$, dann ergibt sich die Übertragungsfunktion des offenen Regelkreises

$$G_0(s) = k \cdot \frac{(1-0,15 \cdot s)}{s^2} \quad (7.20)$$

Die charakteristische Gleichung des geschlossenen Regelkreises lautet

$$\lambda^2 - 0,15 \cdot k \cdot \lambda + k = 0 \quad (7.21)$$

Dieser Regelkreis ist immer instabil, da in der charakteristischen Gleichung ein Koeffizient negativ ist, so daß die notwendige Stabilitätsbedingung nach Hurwitz verletzt ist.

- b) In einem zweiten Versuch wird die Zeitkonstante T_r nicht auf 0,15 gesetzt. Die Übertragungsfunktion des offenen Regelkreises ist (7.18). Für den Frequenzgang gilt damit

$$F_0(j\omega) = k \cdot \frac{(1-0,15 \cdot j\omega) \cdot (1+T_r \cdot j\omega)}{(j\omega)^2 \cdot (1+0,15 \cdot j\omega)} \quad (7.22)$$

Das Aufspalten von $F_0(j\omega)$ in Real- und Imaginärteil liefert

$$F_0(j\omega) = \frac{k \cdot (0,15^2 \cdot \omega^2 - 0,3 \cdot T_r \cdot \omega^2 - 1)}{\omega^2 \cdot (1+0,15^2 \cdot \omega^2)} - j \cdot \frac{k \cdot (T_r - 0,15^2 \cdot \omega^2 \cdot T_r - 0,3)}{\omega \cdot (1+0,15^2 \cdot \omega^2)} \quad (7.23)$$

Wenn $T_r > 0,3$ ist, hat die Ortskurve demnach qualitativ die in Bild 7.6a gezeigten Verläufe.

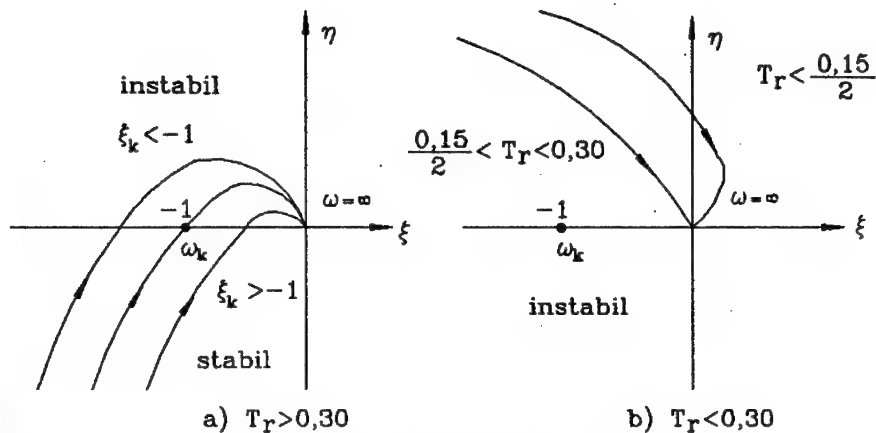


Bild 7.6: Ortskurve des offenen Regelkreises mit PI-Regler.

Wenn $\xi_k < -1$ ist, dann ist der geschlossene Regelkreis immer instabil. Geht die Ortskurve des Frequenzganges durch den Punkt $\xi_k = -1$, so führt der geschlossene Regelkreis stationäre Dauerschwingungen mit der zu ξ_k gehörenden Frequenz ω_k aus. Nur wenn $\xi_k > -1$ ist, ist der geschlossene Regelkreis stabil. Die zu ξ_k gehörende Frequenz ω_k läßt sich durch

$$\eta = -k \cdot \frac{T_r - 0,15^2 \cdot \omega^2 \cdot T_r - 0,3}{\omega \cdot (1 + 0,15^2 \cdot \omega^2)} = 0 \quad (7.24)$$

berechnen. Löst man diese Gleichung auf, dann erhält man

$$\omega_k = \frac{\sqrt{T_r - 0,3}}{0,15 \cdot \sqrt{T_r}} \quad (7.25)$$

Setzt man $\omega = \omega_k$ in

$$\xi = k \cdot \frac{0,15^2 \cdot \omega^2 - 0,3 \cdot T_r \cdot \omega^2 - 1}{\omega^2 \cdot (1 + 0,15^2 \cdot \omega^2)} \quad (7.26)$$

ein, so ergibt sich

$$\xi_k = -0,3 \cdot T_r \cdot k \cdot \frac{0,15^2 + T_r^2 - 0,3 \cdot T_r}{(2 \cdot T_r - 0,3) \cdot (T_r - 0,3)} \quad (7.27)$$

Wenn $\xi_k > -1$ ist, muß

$$k < \frac{(2 \cdot T_r - 0,3) \cdot (T_r - 0,3)}{0,3 \cdot T_r \cdot (0,15^2 + T_r^2 - 0,3 \cdot T_r)} \quad (2.28a)$$

bzw.

$$k_r < \frac{400 \cdot (2 \cdot T_r - 0,3) \cdot (T_r - 0,3)}{0,3 \cdot T_r \cdot k_4 \cdot (0,15^2 + T_r^2 - 0,3 \cdot T_r)} \quad (2.28b)$$

sein. In diesem Fall ist der geschlossene Regelkreis stabil.

Wenn $0,15/2 < T_r < 0,3$ ist, nimmt die zu (7.23) gehörende Ortskurve die in Bild 7.6b gezeigte Form 1 an. Die Form 2 in Bild 7.6b stellt die Ortskurve von $T_r < 0,15/2$ dar. Damit ist aber mit dem Stabilitätskriterium von Nyquist der geschlossene Regelkreis in beiden Fällen instabil. Die Stabilitätsbedingung des Regelkreises mit PI-Regler ist dann

$$\begin{cases} T_r > 0,3 \\ k_r < \frac{400 \cdot (2 \cdot T_r - 0,3) \cdot (T_r - 0,3)}{0,3 \cdot T_r \cdot k_4 \cdot (0,15^2 + T_r^2 - 0,3 \cdot T_r)} \end{cases} \quad (7.29)$$

Wenn eine sprungartige Störung $Z(s) = 1/s$ auftritt, muß der Endwert des Istwertes $y'_a(t)$ Null sein, damit der Regelkreis keine stationäre Regelabweichung hat. Es ist

$$\begin{aligned} Y'_a(s) &= G_z(s) \cdot Z(s) \\ &= \frac{k_4 \cdot (1 + 0,15 \cdot s) \cdot s}{0,15 \cdot s^3 + (1 - 0,15 \cdot T_r \cdot k) \cdot s^2 + (T_r - 0,15) \cdot k \cdot s + k} \end{aligned} \quad (7.30)$$

Mit Hilfe des Endwertsatzes der Laplace-Transformation folgt sofort

$$\lim_{t \rightarrow \infty} y'_a(t) = \lim_{s \rightarrow 0} s \cdot Y'_a(s) = 0 \quad (7.31)$$

Die Festlegung der Parameter k_r und T_r des Reglers spielt die entscheidende Rolle für die Dynamik des Regelkreises. Um das Problem zu vereinfachen, wird die Zeitkonstante T_r zuerst auf 0,4 gesetzt. Dann ist nur die Verstärkung des Reglers k_r festzulegen. Um den konkreten Wert von k_r zu bekommen, wird das Verfahren der Minimierung nach dem linearen Integralkriterium eingesetzt [43]. Diese Methode steht in engem Zusammenhang mit den parameteroptimalen Reglern. Für den Regelkreis zur Erhaltung der besten Fokussierung muß der Istwert y'_a , also das Differential der Fokussierungsfunktion, nach dem Auftreten der Störung schnell wieder zu 0 gehen. Man geht dabei so vor, daß das Integral über $y'_a(t)$, also

$$J(k) = \int_0^{\infty} (y'_a)^2 \cdot dt \quad (7.32)$$

minimiert wird. Zieht man nun den Faltungssatz der Laplace-Transformation heraus, dann kann (7.32) auch

$$J(k) = \frac{1}{2\pi \cdot j} \int_{-j\infty}^{+j\infty} Y'_a(s) \cdot Y'_a(-s) \cdot ds \quad (7.33)$$

geschrieben werden. Dabei wird angenommen, daß die Störung ein Einheitssprung ist. Man kann dann $Y'_a(s)$ und $T_r = 0,4$ in (7.33) einsetzen. Das Integral $J(k)$ läßt sich berechnen. Es lautet

$$J(k) = \frac{k_4^2}{2} \cdot \frac{0,15 \cdot 0,25 \cdot k + 1}{0,1 \cdot k - 0,06 \cdot 0,25 \cdot k^2} \quad (7.34)$$

Es wird nun $J(k)$ nach dem Parameter k minimiert, d.h. aus den notwendigen Bedingungen $dJ(k)/dk = 0$ folgt die Bestimmungsgleichung

$$\frac{k_4^2}{2} \cdot \frac{0,0005625 \cdot k^2 + 0,03 \cdot k - 0,1}{(0,1 \cdot k - 0,015 \cdot k^2)^2} = 0 \quad (7.35)$$

Löst man diese Gleichung auf, so ergibt sich $k = 3,15$ oder

$$k_r = \frac{400 \cdot 3,15}{k_4} \quad (7.36)$$

Die Parameter k_r und T_r können die Stabilitätsbedingung (7.29) erfüllen. Für $Y'_a(s)$ folgt dann

$$Y'_a(s) = \frac{(s^2 + 6,67 \cdot s) \cdot k_4}{s^3 + 5,41 \cdot s^2 + 5,25 \cdot s + 21} \quad (7.37)$$

Als Ergebnis des Reglerentwurfs erhält man für den geschlossenen Regelkreis eine Übertragungsfunktion 3. Ordnung. Um das dynamische Störverhalten des Regelkreises zu bewerten, wird $Y'_a(s)$ zuerst in den Zeitbereich transformiert. Weil die charakteristische Gleichung 3. Ordnung ist, ist es schwer, die Eigenwerte der charakteristischen Gleichung zu bekommen. Deshalb wird $y'_a(t)$ hier durch digitale Simulation berechnet. Die digitale Simulation besteht darin, mit Hilfe eines Rechners für ein gegebenes mathematisches Systemmodell in Form gewöhnlicher Zustandsgleichungen

$$\begin{cases} \dot{x}_i = f_i(x_1, x_2, \dots, x_n; u) \\ i = 1, 2, \dots, n \end{cases} \quad (7.38)$$

und ausgehend von bekannten Anfangswerten $x_i(t_0)$ und für eine vorgegebene Eingangsgröße $u(t)$ über einen Simulationszeitraum \hat{T} , wobei $t_0 < t < t_0 + \hat{T}$ gilt, die zeitlichen Verläufe der n Zustandsvariablen $x_i(t)$ zu ermitteln. Es ist

$$Y'_a(s) = G_Z(s) \cdot Z(s) = \frac{(s^3 + 6,67 \cdot s^2) \cdot k_4}{s^3 + 5,41 \cdot s^2 + 5,25 \cdot s + 21} \cdot Z(s) \quad (7.39)$$

Die zugehörige Differentialgleichung lautet

$$\begin{cases} \ddot{y}'_a(t) + 5,41 \cdot \dot{y}'_a(t) + 5,25 \cdot y'_a(t) + 21 \cdot y'_a(t) = k_4 \cdot \ddot{z}(t) + 6,67 \cdot k_4 \cdot \dot{z}(t) \\ \ddot{y}'_a(0) = \dot{y}'_a(0) = y'_a(0) = 0 \end{cases} \quad (7.40)$$

Man beginnt mit einer Umformung von Gleichung (7.40) in Zustandsform, wobei man als Zustandsvariable definiert:

$$\begin{cases} x_1 = y'_a(t) - k_4 \cdot z(t) \quad , \\ x_2 = 5,41 \cdot y'_a(t) + \ddot{y}'_a(t) - 6,67 \cdot k_4 \cdot z(t) - k_4 \cdot \dot{z}(t) \quad , \\ x_3 = 5,25 \cdot y'_a(t) + 5,41 \cdot \ddot{y}'_a(t) + \ddot{\ddot{y}}'_a(t) - 6,67 \cdot k_4 \cdot \dot{z}(t) - k_4 \cdot \ddot{z}(t) \quad . \end{cases} \quad (7.41)$$

Das der Gleichung (7.40) äquivalente Differentialgleichungssystem 3. Ordnung lautet

$$\begin{cases} \dot{x}_1 = -5,41 \cdot x_1 + x_2 + 1,26 \cdot k_4 \cdot z(t) \quad , \\ \dot{x}_2 = -5,25 \cdot x_1 + x_3 - 5,25 \cdot k_4 \cdot z(t) \quad , \\ \dot{x}_3 = -21 \cdot x_1 - 21 \cdot k_4 \cdot z(t) \end{cases} \quad (7.42)$$

mit $x_1(0) = x_2(0) = x_3(0) = 0$. Der Istwert $y'_a(t)$ ist dann

$$y'_a(t) = x_1(t) + k_4 \cdot z(t) \quad . \quad (7.43)$$

Zur Simulation wird das bekannte Runge-Kutta-Verfahren 4. Ordnung verwendet. Der Algorithmus in der am häufigsten verwendeten Version aus [44] ist:

$$x_{ik+1} = x_{ik} + \frac{h}{6} \cdot (k_{i1} + 2 \cdot k_{i2} + 2 \cdot k_{i3} + k_{i4}) \quad (7.44)$$

mit

$$\begin{aligned} k_{i1} &= f_i(x_{1k}, x_{2k}, \dots, x_{nk}, u_k) \quad , \\ k_{i2} &= f_i(x_{1k} + h \cdot k_{i1}/2, x_{2k} + h \cdot k_{i1}/2, \dots, x_{nk} + h \cdot k_{i1}/2, u_{k+\frac{1}{2}}) \quad , \\ k_{i3} &= f_i(x_{1k} + h \cdot k_{i2}/2, x_{2k} + h \cdot k_{i2}/2, \dots, x_{nk} + h \cdot k_{i2}/2, u_{k+\frac{1}{2}}) \quad , \\ k_{i4} &= f_i(x_{1k} + h \cdot k_{i3}, x_{2k} + h \cdot k_{i3}, \dots, x_{nk} + h \cdot k_{i3}, u_{k+1}) \quad . \end{aligned}$$

Dabei ist

$$\begin{cases} x_k = x(t_k) = x(t_0 + k \cdot h) \\ k = 0, 1, \dots, K \text{ mit } K = \hat{T}/h \end{cases} \quad (7.45)$$

$h = t_{k+1} - t_k$ ist der Zeitschritt der Simulation. In der vorliegenden Arbeit werden $\hat{T} = 5(s)$ und $h = 0,001(s)$ ausgewählt. Das Ergebnis der Simulation wird in Bild 7.7 gezeigt. In diesem Abschnitt wird k_4 auf 1 gesetzt.

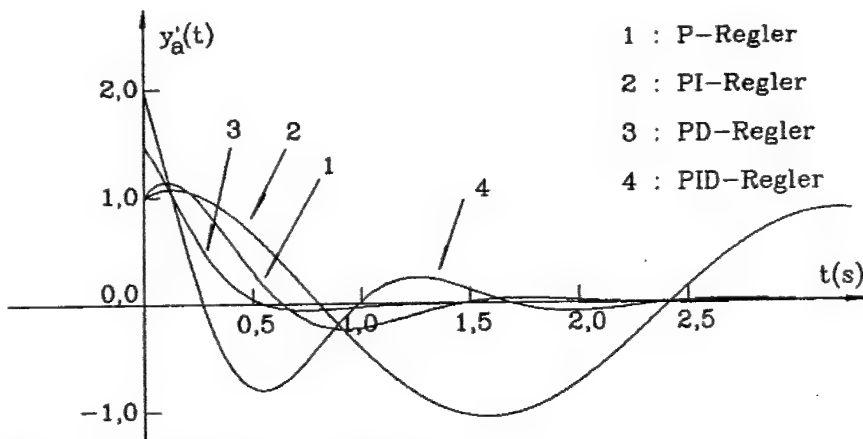


Bild 7.7: Zeitverhalten des Regelkreises.

Die Kurve 2 in Bild 7.7 zeigt, daß, wenn eine sprungartige Störung $z(t)$ auftritt, der Istwert $y'_a(t)$ langsam gedämpft wird. Deshalb benötigt der Regelkreis sehr lange Zeit, um den Istwert $y'_a(t) = 0$ zu erreichen. Da der Vorwärtsweg des Regelkreises ein I-Glied enthält, so darf der Regler im allgemeinen nicht auch noch einen I-Anteil haben. Für den Regelkreis in Bild 7.4 ist der PI-Regler deshalb ungeeignet.

2) PD-Regler

Wenn die Übertragungsfunktion des offenen Regelkreises ein I-Glied enthält, pflegt man einen PD-Regler einzuführen. Der D-Anteil ist bei Änderung der Störgröße wirksam. Er bewirkt, daß das Stellsignal des Reglers nicht nur abhängig von der Regeldifferenz ist, sondern auch von deren Änderungsgeschwindigkeit. Der PD-Regler ist in seiner idealen Form durch die Übertragungsfunktion

$$G_r(s) = k_r \cdot (1 + T_r \cdot s) \quad (7.46)$$

charakterisiert. In diesem Fall lautet die Übertragungsfunktion des offenen Regelkreises

$$G_0(s) = k \cdot \frac{(1 - 0,15 \cdot s) \cdot (1 + T_r \cdot s)}{s \cdot (1 + 0,15 \cdot s)} \quad (7.47)$$

mit $k = (k_r \cdot k_4)/400$. Der PD-Regler kann die Streckenzeitkonstante 0,15 kompensieren. Wenn $T_r = 0,15$ ist, folgt sofort

$$G_0(s) = k \cdot \frac{(1 - 0,15 \cdot s)}{s} \quad (7.48)$$

Die Störübertragungsfunktion des geschlossenen Regelkreises lautet dann

$$G_z(s) = \frac{k_4 \cdot s}{(1 - 0,15 \cdot k) \cdot s + k} \quad (7.49)$$

In diesem Fall lautet die Stabilitätsbedingung

$$k < \frac{1}{0,15} \text{ bzw. } k_r < \frac{400}{0,15 \cdot k_4} \quad (7.50)$$

Die Laplace-Transformierte des Istwertes für $Z(s) = 1/s$ ist

$$Y_a'(s) = \frac{k_4}{(1 - 0,15 \cdot k) \cdot s + k} \quad (7.51)$$

Der Parameter k bzw. k_r wird wieder mit Hilfe der Parameteroptimierung mittels des Güte-Index (7.33) ermittelt. Setzt man (7.51) in (7.33) ein und führt die Integration durch, so erhält man

$$J(k) = \frac{k_4^2}{2 \cdot (1 - 0,15 \cdot k) \cdot k} \quad (7.52)$$

Aus der notwendigen Bedingung

$$\frac{dJ(k)}{dk} = \frac{k_4^2}{2} \cdot \frac{-1 + 0,3 \cdot k}{(1 - 0,15 \cdot k)^2 \cdot k^2} = 0 \quad (7.53)$$

erhält man für den Parameter k den Wert $1/0,3$. Der Reglerparameter k_r ist also

$$k_r = \frac{400}{0,3 \cdot k_4} \quad (7.54)$$

Wird k in (7.51) eingesetzt, so erhält man

$$Y'_a(s) = \frac{2 \cdot k_4}{s + 6} \quad (7.55)$$

Durch Rücktransformation von $Y'_a(s)$ folgt

$$y'_a(t) = 2 \cdot k_4 \cdot \exp(-6 \cdot t) \quad (7.56)$$

Der Istwert $y'_a(t)$ hat in diesem Fall einen exponentiellen Verlauf ohne Überschwingungen.

Nun wird der Fall T_r ungleich $0,15$ untersucht. So kann der Regler die Streckenzeitkonstante $0,15$ nicht kompensieren. Um die Stabilitätsbedingung festzulegen, wird das Nyquist-Kriterium erneut benutzt. Die Übertragungsfunktion des offenen Kreises in diesem Fall lautet

$$F_0(s) = k \cdot \frac{(1 - 0,15 \cdot s) \cdot (1 + T_r \cdot s)}{(1 + 0,15 \cdot s) \cdot s} \quad (7.57)$$

und für den Frequenzgang gilt

$$\begin{aligned} F_0(j\omega) = & k \cdot \frac{0,15 \cdot (1 + 0,15 \cdot T_r \cdot \omega^2) - T_r + 0,15}{(1 + 0,15^2 \cdot \omega^2)} \\ & - j \cdot k \cdot \frac{1 + 0,15 \cdot T_r \cdot \omega^2 + 0,15 \cdot \omega^2 \cdot (T_r - 0,15)}{\omega \cdot (1 + 0,15^2 \cdot \omega^2)} \quad (7.58) \end{aligned}$$

- a) Wählt man die Zeitkonstante $T_r > 0,15/2$ aus, dann nimmt die zu (7.58) gehörende Ortskurve die in Bild 7.8a gezeigte Form an.

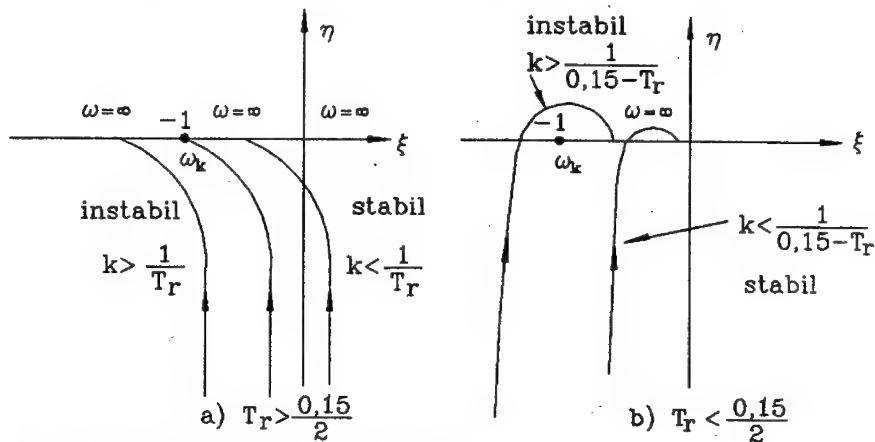


Bild 7.8: Ortskurve mit einem PD-Regler

Nur wenn die Bedingung

$$k < \frac{1}{T_r} \text{ bzw. } k_r < \frac{400}{k_4 \cdot T_r} \quad (7.59)$$

erfüllt ist, ist der Regelkreis stabil.

- b) Für $T_r < 0,15/2$ hat die Ortskurve die in Bild 7.8b gezeigte Gestalt. Die Frequenz ω_k läßt sich durch

$$\eta = -k \cdot \frac{1 - \omega^2 \cdot (0,15^2 - 0,3 \cdot T_r)}{\omega \cdot (1 + 0,15^2 \cdot \omega^2)} = 0 \quad (7.60)$$

berechnen. Es ist $\omega_k = 1/\sqrt{0,15^2 - 0,3 \cdot T_r}$. Bei dieser Frequenz ist

$$\xi_k = -k \cdot (0,15 - T_r) \quad (7.61)$$

Wenn die Bedingung $\xi_k > -1$, also

$$k < \frac{1}{0,15 - T_r} \quad (7.62)$$

erfüllt ist, ist das System stabil. Nun wird $Z(s) = 1/s$ gesetzt. Dann lautet die Laplace-Transformierte des Istwertes

$$Y'_a(s) = G_z(s) \cdot Z(s) = k_4 \cdot \frac{0,15 \cdot s + 1}{0,15 \cdot (1 - T_r \cdot k) \cdot s^2 + (1 - 0,15 \cdot k + T_r \cdot k) \cdot s + k} \quad (7.63)$$

Für $T_r = 0,1$ lautet das Integral (7.33)

$$J(k) = \frac{k_4^2}{2} \cdot \frac{1 + 0,05 \cdot k}{k - 0,15 \cdot k^2 + 0,005 \cdot k^3} \quad (7.64)$$

Löst man die Gleichung

$$\frac{dJ(k)}{dk} = \frac{k}{2} \cdot \frac{-0,0005 \cdot k^3 - 0,0075 \cdot k^2 + 0,3 \cdot k - 1}{(k - 0,15 \cdot k^2 + 0,005 \cdot k^3)^2} = 0 \quad (7.65)$$

nach k auf, erhält man für k den Wert 3,78. Die Verstärkung k kann die Stabilitätsbedingung (7.59) erfüllen. Setzt man k und T_r in $Y'_a(s)$ ein, so folgt

$$Y'_a(s) = k_4 \cdot \frac{0,15 \cdot s + 1}{0,093 \cdot s^2 + 0,811 \cdot s + 3,78} \quad (7.66)$$

Führt man nun die Rücktransformation aus, so bekommt man

$$y'_a(t) = 1,80 \cdot k_4 \cdot \exp(-4,36 \cdot t) \cdot \sin(4,642 \cdot t + 1,109). \quad (7.67)$$

Nach Bild 7.4 kann die Kameraposition $p'(t)$ berechnet werden. Es ist

$$p'(t) = y'_a(t)/k_4 - z(t) \quad (7.68a)$$

Setzt man $k_4 = 1$ und $z(t) = 1$ in $p'(t)$ ein, so ergibt sich

$$p'(t) = -1 + 1,80 \cdot \exp(-4,36 \cdot t) \cdot \sin(4,642 \cdot t + 1,109) \quad (7.68b)$$

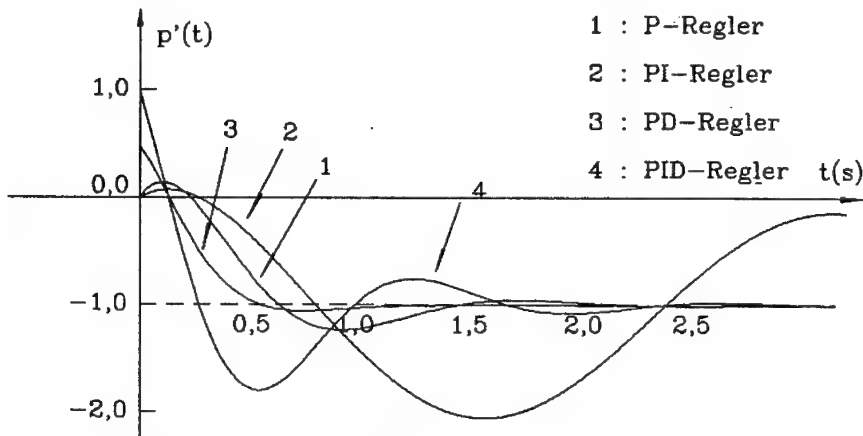


Bild 7.9: Verlauf der Funktion $p'(t)$.

Die Kurve 3 in Bild 7.7 zeigt die Störantwort $y'_a(t)$ von (7.67). An diesem Ergebnis erkennt man zweierlei. Zum einen ist die Dämpfung des Regelkreises gegenüber dem Regelkreis mit PI-Regler erhöht. Das System braucht nur kurze Zeit, um den Istwert $y'_a(t) = 0$ zu erreichen. Zum anderen verursacht der D-Anteil des Reglers den starken Stelleingriff zu Beginn des Regelvorgangs. Deshalb wird die Kamera am Anfang des Regelvorgangs eine sprunghafte Bewegung durchführen (Kurve 3 in Bild 7.9). Durch diese Bewegung ist die Kamera weiter von der besten Fokussierungsposition entfernt. Das Ergebnis ist unerwünscht.

3) PID-Regler

Nun erfolgt die weitere Untersuchung des gleichen Problems mit einem PID-Regler. Der PID-Regler eignet sich für besonders schwierige und schnelle Regelungen. Er vereinigt die rasche Wirkung des P-Reglers, die stabilisierende Wirkung des D-Anteils und den I-Anteil, der solange eingreift, bis die Regelabweichung Null wird. Die Übertragungsfunktion des PID-Reglers gilt im idealen Fall

$$G_r(s) = k_r \cdot \frac{(1+T_{r1} \cdot s) \cdot (1+T_{r2} \cdot s)}{s} \quad (7.69)$$

Jetzt sind die freien Reglerparameter k_r , T_{r1} , T_{r2} zu wählen. Die Übertragungsfunktion des offenen Regelkreises $G_0(s)$ lautet

$$G_0(s) = k \cdot \frac{(1-0,15 \cdot s) \cdot (1+T_{r1} \cdot s) \cdot (1+T_{r2} \cdot s)}{s^2 \cdot (1+0,15 \cdot s)} \quad (7.70)$$

Dabei ist k gleich $k_r \cdot k_4 / 400$. Es wird im Sinne der Kompensation gesetzt

$$T_{r1} = 0,15 \quad (7.71)$$

und (7.70) reduziert sich damit auf

$$G_0(s) = k \cdot \frac{(1-0,15 \cdot s) \cdot (1+T_{r2} \cdot s)}{s^2} \quad (7.72)$$

Die charakteristische Gleichung des geschlossenen Regelkreises ist

$$(1-0,15 \cdot k \cdot T_{r2}) \cdot \lambda^2 + (T_{r2} - 0,15) \cdot k \cdot \lambda + k = 0 \quad (7.73)$$

Die notwendigen und hinreichenden Stabilitätsbedingungen für den Regelkreis sind dann

$$\begin{cases} T_{r2} > 0,15 \\ 0 < k_r < \frac{400}{0,15 \cdot T_{r2} \cdot k_4} \end{cases} \quad (7.74)$$

Die Parameter k_r und T_{r2} werden erneut mit Hilfe des Verfahrens der Parameteroptimierung bestimmt. Dazu muß die Laplace-Transformierte des Istwertes $Y_a'(s)$ bekannt sein. Wenn $Z(s) = 1/s$ ist, folgt

$$\begin{aligned} Y_a'(s) &= G_z(s) \cdot Z(s) \\ &= \frac{k_4 \cdot s}{(1-0,15 \cdot T_{r2} \cdot k) \cdot s^2 + (T_{r2} - 0,15) \cdot k \cdot s + k} \end{aligned} \quad (7.75)$$

Das Integral (7.33) ergibt sich nach kurzer Rechnung

$$J(k, T_{r2}) = \frac{k_4^2}{2 \cdot k \cdot (T_{r2} - 0,15) \cdot (1 - 0,15 \cdot T_{r2} \cdot k)} \quad (7.76)$$

Die optimalen Parameter k und T_{r2} können aus dem Gleichungssystem

$$\begin{cases} \frac{\partial J(k, T_{r2})}{\partial k} = \frac{k_4^2 \cdot T_{r2} + 0,15 + 0,3 \cdot T_{r2}^2 \cdot k - 2 \cdot 0,15^2 \cdot T_{r2} \cdot k}{[k \cdot (T_{r2} - 0,15) \cdot (1 - 0,15 \cdot T_{r2} \cdot k)]^2} = 0 \\ \frac{\partial J(k, T_{r2})}{\partial T_{r2}} = \frac{k_4 \cdot (-k + 0,3 \cdot T_{r2} \cdot k^2 - 0,15^2 \cdot k^2)}{[k \cdot (T_{r2} - 0,15) \cdot (1 - 0,15 \cdot T_{r2} \cdot k)]^2} = 0 \end{cases} \quad (7.77)$$

berechnet werden. Nach einfacher Rechnung ergibt sich

$$\begin{cases} k = 1/0,15^2 \\ T_{r2} = 0,15 \end{cases} \quad (7.78)$$

Das Ergebnis ist unerwünscht, da der Regelkreis sich in dem kritischen Stabilitätszustand befindet. Setzt man k und T_{r2} in $Y'_a(s)$ ein, dann erhält man

$$Y'_a(s) = 0,15^2 \cdot k_4 \cdot s \quad (7.79)$$

Die Rücktransformation von $Y'_a(s)$ ergibt

$$y'_a(t) = 0,15^2 \cdot k_4 \cdot \delta(t) \quad (7.80)$$

Das ist in der Praxis nicht realisierbar. In einem weiteren Versuch wird die Zeitkonstante T_{r2} auf 0,3 gesetzt. Mit Hilfe des Verfahrens der Parameteroptimierung erhält man sofort

$$k = \frac{1}{0,3^2} \quad (7.81)$$

Die Parameter k und T_{r2} können auch die Stabilitätsbedingung (7.74) erfüllen. Die Laplace-Transformierte des Istwertes lautet in diesem Fall

$$Y'_a(s) = \frac{2 \cdot k_4 \cdot s}{s^2 + 3,333 \cdot s + 22,222} \quad (7.82)$$

Durch Rücktransformation in den Zeitbereich erhält man unmittelbar

$$y'_a(t) = 2,138 \cdot k_4 \cdot \exp(-1,667 \cdot t) \cdot \cos(4,41 \cdot t + 0,362) \quad (7.83)$$

Die Kurve 4 in Bild 7.7 zeigt den Verlauf von (7.83). Weil der PID-Regler auch einen I-Anteil enthält, wird der Regelkreis noch nicht genügend gedämpft. Außerdem verursacht der D-Anteil des Reglers den starken Stelleingriff zu Beginn des Regelvorgangs. Am Anfang des Regelvorgangs führt die Kamera auch eine sprunghafte Bewegung durch (Kurve 4 in Bild 7.9).

4) P-Regler

Schließlich wird untersucht, die Strecke mit einem einfachen P-Regler zu regeln. Der Grund dafür ist, daß der Vorwärtszweig des Regelkreises in Bild 7.4 ein Integrationsglied enthält. Der P-Regler gewährleistet in diesem Fall eine völlige Störungsausregelung, da die Störgröße erst hinter dem I-Glied in der Strecke eingreift. Die Übertragungsfunktion des P-Reglers ist

$$G_r(s) = k_r \quad (7.84)$$

Setzt man (7.84) in (7.16) ein, ergibt sich

$$G_z(s) = \frac{k_4 \cdot s \cdot (1 + 60,15 \cdot s)}{0,15 \cdot s^2 + (1 - 0,15 \cdot k) \cdot s + k} \quad (7.85)$$

Es wird jetzt die Stabilität des Regelkreises untersucht. Die charakteristische Gleichung von (7.85) ist

$$\lambda^2 + \frac{1 - 0,15 \cdot k}{0,15} \cdot \lambda + k = 0 \quad (7.86)$$

Aus der notwendigen und hinreichenden Stabilitätsbedingung folgt

$$\frac{1 - 0,15 \cdot k}{0,15} > 0 \quad (7.87)$$

oder

$$k < \frac{1}{0,15} \text{ bzw. } k_r < \frac{400}{0,15 \cdot k_4} \quad (7.88a)$$

und $k = k_r \cdot k_4 / 400 > 0$, da $k_4 > 0$ ist, muß

$$k_r > 0 \quad (7.88b)$$

sein, was zusammen mit (7.88a) und (7.88b) für k_r den Wertebereich

$$0 < k_r < \frac{400}{0,15 \cdot k_4} \quad (7.88c)$$

ergibt. Wenn eine sprungartige Störung $Z(s) = 1/s$ auftritt, ist es

$$Y'_a(s) = G_Z(s) \cdot Z(s) = \frac{k_4 \cdot (1 + 0,15 \cdot s)}{0,15 \cdot s^2 + (1 - 0,15 \cdot k) \cdot s + k} \quad (7.89)$$

(7.88c) bestimmt den Wertebereich von k_r . Weil der Regler nur einen wählbaren Parameter hat, spielt k_r die entscheidende Rolle für das Zeitverhalten des Regelkreises. Um den konkreten Wert von k_r zu bekommen, wird das Verfahren der Parameteroptimierung nochmal eingesetzt. Das Integral $J(k)$ in diesem Fall lautet

$$J(k_r) = \frac{k_4}{2} \cdot \frac{(1 + 0,15 \cdot k)}{(k - 0,15 \cdot k^2)} \quad (7.90)$$

Die Bestimmungsgleichung für k folgt

$$\frac{dJ(k)}{dk} = \frac{k_4^2}{2} \cdot \frac{0,15^2 \cdot k^2 + 0,3 \cdot k - 1}{(k - 0,15 \cdot k^2)^2} = 0 \quad (7.91)$$

Löst man diese Gleichung auf, so ergibt sich

$$k = \frac{1}{0,36} \text{ bzw. } k_r = \frac{400}{0,36 \cdot k_4} \quad (7.92)$$

Der Parameter k_r kann die Stabilitätsbedingung (7.88c) erfüllen. Setzt man $k_r = 1/0,36$ in $Y'_a(s)$ ein, so folgt

$$Y'_a(s) = \frac{k_4 \cdot (0,15 \cdot s + 1)}{0,15 \cdot s^2 + 0,583 \cdot s + 2,778} \quad (7.93)$$

Die Systemantwort auf $z(t) = \sigma(t)$ ist dann

$$y'_a(t) = 1,585 \cdot k_4 \cdot \exp(-1,95 \cdot t) \cdot \sin(3,84 \cdot t + 0,683) \quad (7.94)$$

Kurve 1 in Bild 7.7 ist der Verlauf von (7.94). Einerseits ist der Regelkreis genügend gedämpft, andererseits hat die Kamera zu Beginn des Regelvorgangs keine sprungförmige Bewegung durchgeführt. Außerdem ist der P-Regler in der Praxis einfach zu realisieren. Deshalb wird der P-Regler im Regelkreis zur Erhaltung der besten Fokussierungsposition verwendet.

7.5 Zeitverhalten

In diesem Abschnitt wird die Varianz der Grauwerte als Fokussierungsfunktion benutzt. Bild 7.10, Bild 7.11 und Bild 7.12 zeigen das Zeitverhalten des Regelkreises zur Erreichung der besten Fokussierung durch Experimente. In den Experimenten ist das Objekt unbeweglich und die beste Fokussierungsposition liegt in der Position $p' = 150$ mm. Die Zeit für jeden Schritt umfaßt die Zeit für die Berechnung der Fokussierungsfunktion und die Zeit für die Bewegung der Kamera. Die max. Bewegungsgeschwindigkeit der Kamera beträgt 20 mm/Sekunde. Die Zeit für die Berechnung der Varianz dauert 0,24 Sekunden. Wie in Abschnitt 7.2 geschildert, wird die Schrittgröße der Kamera vor dem Fokussierungsprozeß bestimmt; deshalb ist die Konstante ΔF der einzige Faktor zur Beeinflussung des Zeitverhaltens des Regelkreises.

Bild 7.10a und b sind der Verlauf der Varianz $VAR(t)$ und Verlauf der Kameraposition $p'(t)$ für die Konstante ΔF gleich 1. In diesem Fall ist die beste Fokussierungsposition problemlos gefunden.

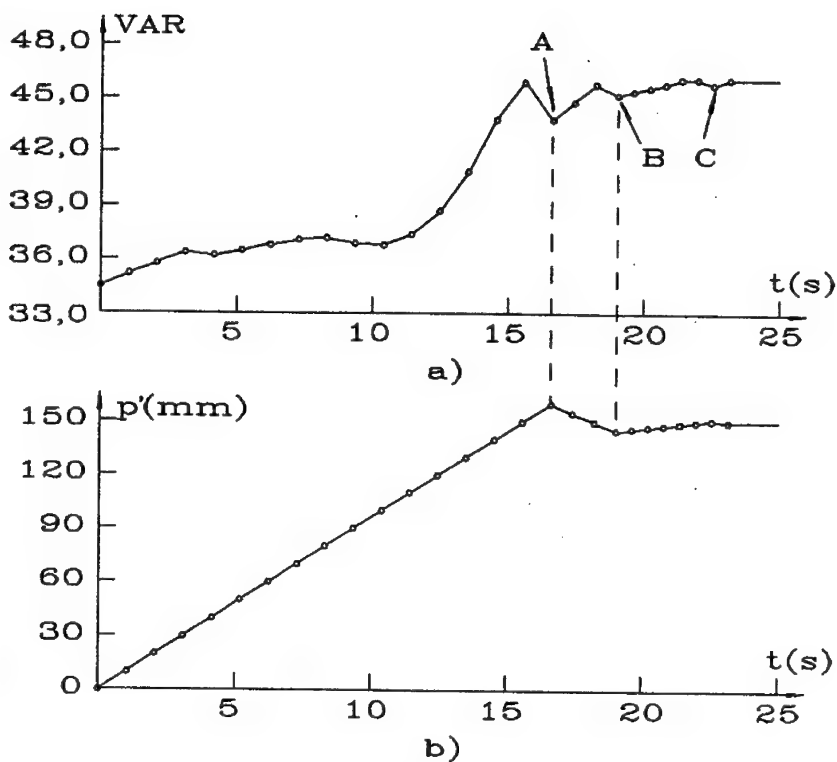
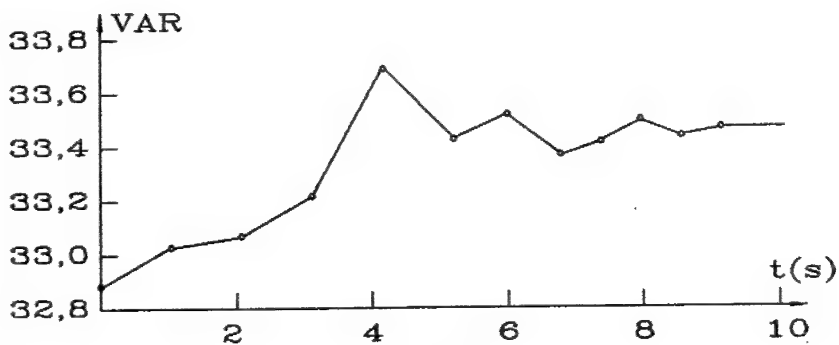


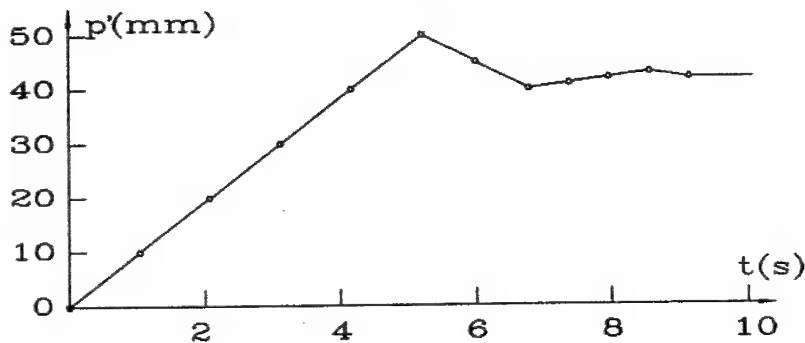
Bild 7.10: a) Verlauf der Zeitfunktion $VAR(t)$,
b) Verlauf der Zeitfunktion $p'(t)$.

In Bild 7.10a befindet sich die Kamera in den ersten 7 Sekunden im unwirksamen Bereich. Die Varianz nimmt in diesem Bereich nicht monoton zu, sie verändert sich nur wenig. Es gibt keinen Fall, in dem die Bedingung (7.5) erfüllt wird. Die Kamera wird bis zum Punkt A, wo die Bedingung (7.5) erfüllt ist, fahren. Im Punkt A hat die Kamera die beste Fokussierungsposition überschritten. Deshalb muß die Kamera zurückfahren. In der zweiten Phase fährt die Kamera mit der Schrittgröße 5 mm bis zum Punkt B, wo die Bedingung (7.4) erfüllt ist. In der letzten Phase fährt die Kamera in umgekehrter Richtung mit der Schrittgröße 1 mm. Wenn die Bedingung (7.4) wieder erfüllt ist, fährt die Kamera 1 mm zurück. Die Position $p' = 150$ mm ist der stationäre Endwert des Istwertes vom Regelkreis.

Wenn die Konstante ΔF zu klein ist, wird die Kamera eine falsche Position als beste Fokussierungsposition finden. Bild 7.11 zeigt das Zeitverhalten des Regelkreises für $\Delta F = 0,1$. Wegen des Einflusses des Bildrauschens hat die Varianz in der Position $p' = 42 \text{ mm}$ ein lokales Maximum. In der Position $p' = 50 \text{ mm}$ ist $FF_n - FF_{n-1} = -0,27 < -0,1$. Deshalb ist der Computer davon überzeugt, daß die Kamera die beste Fokussierungsposition überschritten hat. Am Ende des Fokussierungsprozesses wird die Position des lokalen Maximums als die beste Fokussierungsposition betrachtet. Dieser Regelkreis ist zwar stabil, sein stationärer Endwert ist aber falsch.



a)



b)

Bild 7.11: a) Verlauf der Zeitfunktion $VAR(t)$,
b) Verlauf der Zeitfunktion $p'(t)$.

Wenn die Konstante ΔF zu groß ist, ist der Regelkreis instabil. Das kann man an Bild 7.12 erkennen.

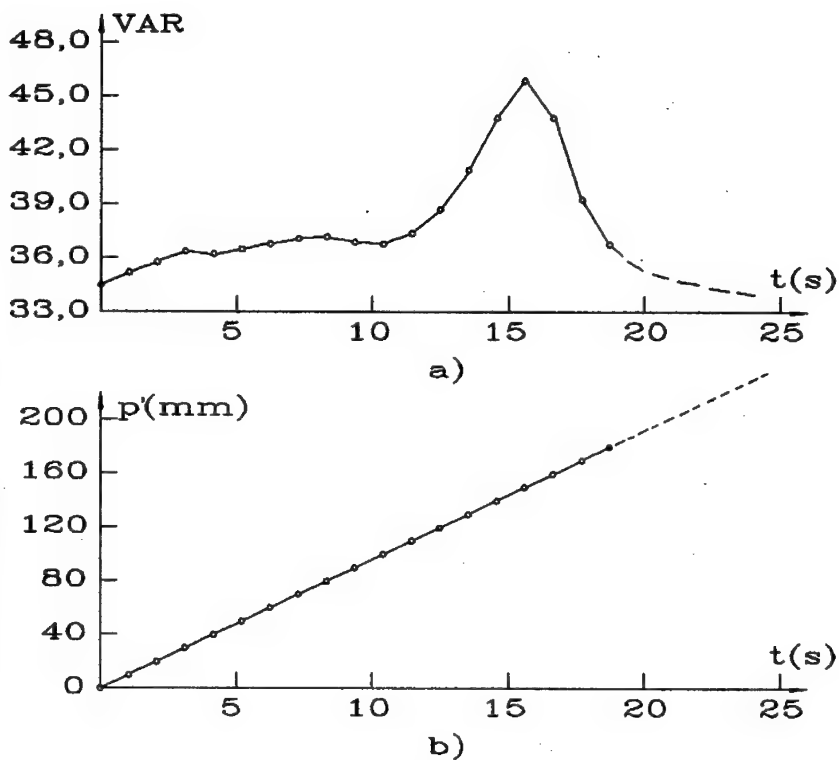


Bild 7.12: a) Verlauf der Zeitfunktion $VAR(t)$,
b) Verlauf der Zeitfunktion $p'(t)$.

Für $\Delta F = 3$ gibt es im ganzen Arbeitsbereich keine Position, wo die Bedingung $FF_n - FF_{n-1} < -3$ erfüllt werden kann. Die Kamera wird bis zum Rand des Arbeitsbereiches fahren. Deshalb ist der Regelkreis instabil.

Das Zeitverhalten vom Regelkreis zur Erhaltung der besten Fokussierung wird zuerst theoretisch analysiert. Dasselbe Zeitverhalten wird auch experimentell gemessen. In dem Experiment wird das Muster Honig als Objekt benutzt. Die Konstante k_4 vom Muster Honig ist ca. 0,115. Wie es im Abschnitt 7.4 geschildert ist, wird ein P-Regler im Regelkreis verwendet. Setzt man $k_4 = 0,115$ und $k = 1/0,36$ in (7.85) ein, so folgt

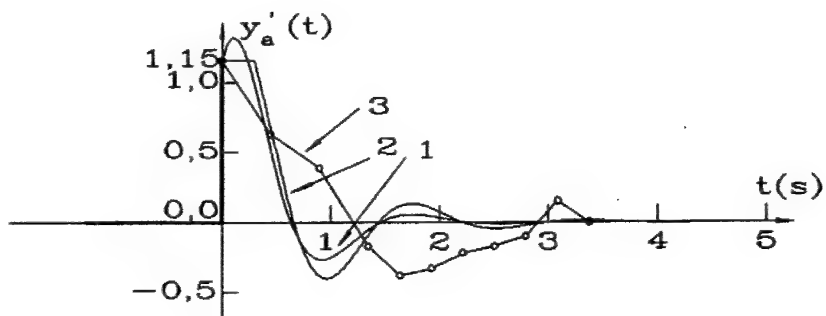
$$G_z(s) = \frac{0,115 \cdot s \cdot (0,15 \cdot s + 1)}{9,15 \cdot s^2 + 0,583 \cdot s + 2,777} \quad (7.95)$$

Wenn die Störgröße $z(t) = 10 \cdot \sigma(t)$ ist, also $Z(s) = 10/s$, dann erhält man

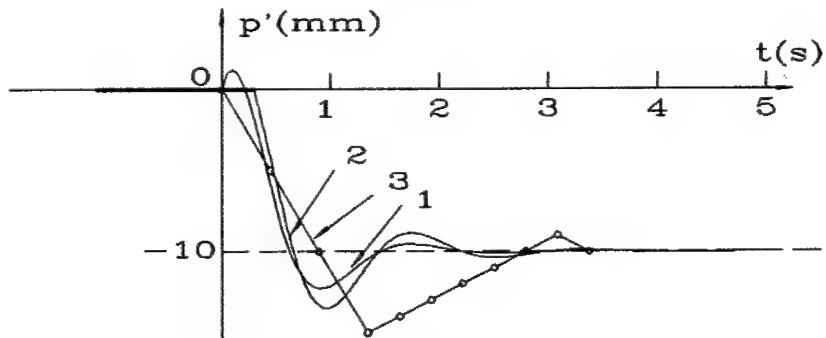
$$Y'_a(s) = G_z(s) \cdot Z(s) = \frac{1,15 \cdot (s + 6,67)}{s^2 + 3,89 \cdot s + 18,52} \quad (7.96)$$

Die Rücktransformation von $Y'_a(s)$ ergibt

$$y'_a(t) = 1,823 \cdot \exp(-1,95 \cdot t) \cdot \sin(3,84 \cdot t + 0,683) \quad (7.97)$$



a)



b)

Bild 7.13: a) Funktion $y'_a(t)$.
b) Funktion $p'(t)$.

Die Laplace-Transformation der Kameraposition $p'(t)$ läßt sich aus Bild 7.4 berechnen. Es ist

$$P'(s) = \frac{2,78 \cdot s - 18,52}{s^2 + 3,89 \cdot s + 18,52} \cdot \frac{1}{s} \quad (7.98)$$

Der Verlauf von $p'(t)$ ist dann

$$p'(t) = -10 + 15,85 \cdot \exp(-1,95 \cdot t) \cdot \sin(3,84 \cdot t + 0,836) \quad (7.99)$$

Die Funktionen von $y'_a(t)$ und $p'(t)$ werden in Bild 7.13 (Kurve 1) dargestellt. Zu Beginn des Regelvorgangs wird der Istwert $y'_a(t)$ weiter zunehmen und die Kamera fährt in die falsche Richtung. Das steht nicht mit dem Ergebnis des realen Regelkreises im Einklang. Der Grund ist, wenn das Totzeitglied durch eine Pade-Approximation erster Ordnung angenähert ist, wird ein Fehler erzeugt. Den Einfluß dieses Fehlers kann man durch den Vergleich der Kurve 1 und Kurve 2 in Bild 7.13 deutlich bemerken. Kurve 2 ist das Zeitverhalten des Regelkreises, wenn das Totzeitglied nicht angenähert wird (exakte Lösung). Kurve 2 ist mit Hilfe der digitalen Simulation des Regelkreises in Bild 7.14 ermittelt worden.

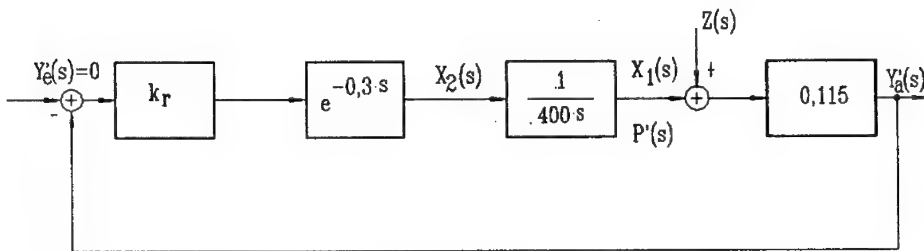


Bild 7.14: Linearer Regelkreis mit Totzeit.

Der Algorithmus der Simulation ist das Runge-Kutta-Verfahren 4. Ordnung. Dafür muß die Differentialgleichung des Regelkreises zuerst in Zustandsform umgeformt werden. Durch Einführen von geeigneten Zustandsvariablen

$$\begin{cases} \dot{x}_1(t) = p'(t) = 8,696 \cdot y'_a(t) - z(t) \\ \dot{x}_2(t) = 400 \cdot \dot{x}_1(t) = -k_r \cdot y'_a(t-0,3) \end{cases} \quad (7.100)$$

kann man das Zustandsgleichungssystem von Bild 7.14 bilden. Es ist

$$\begin{cases} \dot{x}_1(t) = \frac{1}{400} \cdot x_2(t) \\ x_2(t) = -0,115 \cdot k_r \cdot [x_1(t-0,3) + z(t-0,3)] \end{cases} \quad (7.101)$$

mit $x_1(0) = 0$ und $x_2(0) = 0$. Die Funktionen $y'_a(t)$ und $p'(t)$ lassen sich zu

$$\begin{cases} y'_a(t) = 0,115 \cdot [x_1(t) + z(t)] \\ p'(t) = x_1(t) \end{cases} \quad (7.102)$$

bestimmen. Das Ergebnis der Simulation ist auch in Bild 7.13 gezeigt. Da die Übertragungsfunktion der Interfacekarte als Totzeitglied bleibt, bewegt sich die Kamera in den ersten 0,3 Sekunden nicht. Nach 0,3 Sekunden wird die Kamera direkt zur besten Fokussierungsposition fahren.

Die Kurve für $t < 0$ stellt den Zustand der besten Fokussierung dar. Bei der besten Fokussierung liegt die Kamera in der Position $p' = 0$ und der Istwert des Regelkreises ist Null. Zum Zeitpunkt $t = 0$ wird eine Störung $z(t) = 10$ plötzlich aufgegeben. Wegen der Trägheit des Systems kann die Kamera nicht sofort darauf reagieren. Der Istwert ist von 0 auf 1,15 gestiegen. Die Eingabe am Regler ist - 1,15. Unter der Auswirkung von diesem Signal gibt der Regler ein Signal aus. Die Kamera fährt zurück, anschließend wird die Kamera 10 mm zurückgefahren.

Durch die theoretische Analyse kann man bemerken, daß mit Hilfe des einfachen P-Reglers alle Anforderungen am Regelkreis erfüllt werden können. In der Praxis ist der Kontrollalgorithmus des Reglers ähnlich wie beim P-Regler, allerdings ist die Schrittgröße der Kamera veränderbar. Sie ist abhängig vom Istwert y'_a . Am Anfang des Fokussierungsprozesses ist die Kamera normalerweise weit von der besten Fokussierungsposition entfernt. Der Istwert ist deshalb ziemlich groß. Deshalb wird die Kamera mit größerer Schrittgröße gefahren. In der Nähe der besten Fokussierungsposition ist der Istwert $y'_a(t)$ klein. In dieser Situation fährt die Kamera mit kleiner Schrittgröße.

Kurve 3 in Bild 7.13a und b sind die experimentellen Ergebnisse von $y'_a(t)$ und $p'(t)$. Am Anfang fährt die Kamera mit der größeren Schrittgröße 5 mm zurück. Wenn der Istwert $y'_a(t)$ sich verkleinert, wird die Schrittgröße auf 1 mm gesetzt. Der Endwert von $p'(t)$ ist - 10 mm.

Durch den Vergleich der Kurven in Bild 7.13a und b kann man bemerken, daß die Ergebnisse des Zeitverhaltens zwischen der theoretischen Analyse und den experimentellen Ergebnissen sehr ähnlich sind. Es gibt aber auch Unterschiede. Die Ursache dafür ist vielfältig und zwar:

1. Die Übertragungsfunktion des Bildverarbeitungssystems ist mit einer konstanten Funktion näherungsweise dargestellt. Die Konstante k_4 ist von Objekt zu Objekt unterschiedlich.
2. Die Übertragungsfunktion des Schrittmotors ist als Verstärker modelliert. In Wirklichkeit ist der Schrittmotor in seinem Aufbau und seiner Wirkungsweise mit einem bürstenlosen Gleichstrommotor vergleichbar.
3. Die Übertragungsfunktion der Interfacekarte ist ein Totzeitglied. Durch die Rationalisierung wird dieses Totzeitglied als die Pade-Approximation erster Ordnung angenähert.
4. Die Herleitung der Übertragungsfunktion vom Bildverarbeitungssystem hat die Rechenzeit der Fokussierungsfunktion nicht berücksichtigt.

8. Zusammenfassung

Um eine Produktion von Bildern mit höchster Qualität zu bekommen, ist eine eindeutige Definition des Optimalfokus erforderlich. In der vorliegenden Arbeit wird zuerst die Bedeutung der automatischen Fokussierung erläutert und es wird auf die Geschichte der automatischen Fokussierung zurückgeblickt (Kapitel 1).

In Kapitel 2 wird das Fokussierungsprinzip des menschlichen Auges erklärt. Einige für die Fokussierung wichtige Begriffe werden aufgestellt und diskutiert. Die optische Übertragungsfunktion (OÜF) wird von der Wellenoptik und der geometrischen Optik hergeleitet. Um die Bildqualität zu beurteilen, werden insgesamt 4 Kriterien aufgestellt.

Die Grundidee der automatischen Fokussierung für die Kamera ist die Suche nach den Fokussierungsfunktionen. Aus diesem Grund werden einige Anforderungen, die an Fokussierungsfunktionen gestellt werden, vorgestellt. Insgesamt wird das automatische Fokussierungsverfahren mit 8 verschiedenen Fokussierungsfunktionen realisiert (Kapitel 3). Die Fokussierungsprinzipien von ihnen werden in Kapitel 4 genau analysiert.

Zur Beurteilung der Fokussierungsfunktionen werden Experimente durchgeführt und die Kurven der Fokussierungsfunktionen in verschiedenen Situationen gemessen und verglichen. Durch die experimentellen Ergebnisse werden die Vorteile und die Nachteile der verschiedenen Fokussierungsfunktionen herausgearbeitet.

In Kapitel 5 wird das Problem der Fokussierung von Bildausschnitten untersucht und die Grundidee zur Lösung dieses Problems erläutert. Die Auswirkungen des Bildrauschens auf die Fokussierungsfunktionen werden ebenfalls untersucht. Zwei Methoden zur Unterdrückung der Auswirkungen des Bildrauschens werden diskutiert und realisiert. Am Ende von Kapitel 5 wird die Zoom-Auswirkung der Kamera untersucht. Ein Algorithmus zur Kompensierung der Zoom-Auswirkung wird vorgestellt.

Die Fokussierungsfunktion ist ungeeignet zur Fokussierung von beweglichen Objekten. Deshalb wird ein neues Verfahren zur Fokussierung von beweglichen Objekten entwickelt. Das Verfahren zur Fokussierung von beweglichen Objekten wird mit Hilfe eines Lasergerätes realisiert. Die Objektweite des Objektes als

die Funktion der Bildkoordinaten vom Laserfleck wird hergeleitet. Durch die sogenannte Kamerakalibrierung werden die Systemkoeffizienten bestimmt. Der Systemfehler wird durch Experimente und durch eine theoretische Analyse untersucht (Kapitel 6).

Zur Entwicklung der automatischen Fokussierung wird das Bildverarbeitungssystem Max Vision AT-1 benutzt. Um Fokussierungsregelkreise zu entwerfen, wird der Fokussierungsprozeß in zwei Phasen gemäß der gestellten Aufgabenstellung eingeteilt, nämlich die Erreichung der besten Fokussierung und die Erhaltung der besten Fokussierung. Für beide Phasen wird jeweils ein mathematisches Modell der Fokussierungsregelkreise vorgestellt. Die Übertragungsfunktion des Fokussierungsregelkreises für die Erhaltung der besten Fokussierung kann durch Linearisierung des Übertragungsgliedes Bildverarbeitungssystem und Rationalisierung des Totzeitgliedes erhalten werden. Gemäß dieser Übertragungsfunktion wird der Regler des Fokussierungsregelkreises entworfen. Schließlich wird das Zeitverhalten des Systems durch Experimente und eine theoretische Analyse erhalten (Kap. 7).

Ausblick und kritische Würdigung

Aufbauend auf dem derzeitigen Stand der Untersuchungen sollten weiterführende Arbeiten an folgenden Punkten ansetzen.

- 1) Um die Zoom-Auswirkung der Kamera zu vermeiden, sollte die automatische Fokussierung durch die Veränderung der Brennweite der Kamera erreicht werden.
- 2) Das Verfahren der Fokussierung von Bildausschnitten sollte weiter untersucht werden, damit die Tiefeninformation einer Szene aus zweidimensionalen Bildern extrahiert werden kann.
- 3) Der in Kapitel 6 geschilderte Algorithmus zur direkten Berechnung der Objektweite mit Hilfe eines Lasergerätes sollte weiterentwickelt werden, um ein Robotersichtsystem aufzubauen.

In der vorliegenden Arbeit werden verschiedene Verfahren zur automatischen Fokussierung des Objektes untersucht. Durch die theoretische Analyse wird die Frequenzreaktion des defokalen optischen Systems geklärt. Viele interessante Ergebnisse sind durch Experimente erzielt worden.

9. Wichtige Formelzeichen

$SI(w)$	Schärfeindex des Bildes
w	fokale Variable
$f(x,y), g(x,y)$	Grauwert des Bildpunktes in der Position (x,y)
(x,y)	Ortskoordinaten des Bildpunktes
R, L	Dimension in der Bildebene
f	Brennweite der Kamera
p	Objektweite
q	Bildweite
n	Brechungsindex
B	Vergrößerungsfaktor der Kamera
BZ	Blendenzahl der Linse
$N.A.$	numerische Apertur
α	Öffnungswinkel der Linse
$S.T.$	Schärfentiefe der Kamera
$F.T.$	Fokussierungstiefe der Kamera
r	Durchmesser der Linse
Δu	Durchmesser vom Unschärfebereich
$i(x,y)$	Beleuchtungsstärke in der Position (x,y)
$l(x,y)$	Reflexionsgrad in der Position (x,y)
$E = [f(x,y)]$	Bildmatrix des digitalisierten Bildes
$h(x,y)$	Punktantwort
$F(u,v)$	Fouriertransformation vom Bild $f(x,y)$
(u,v)	Ortsfrequenzen
$H(u,v)$	optische Übertragungsfunktion
(ξ', η')	reduzierte Pupillenkoordinaten für Eintrittspupille
(ξ, η)	reduzierte Pupillenkoordinaten für Austrittspupille
$b(\xi, \eta)$	Pupillenfunktion
σ	Ausdehnungsparameter des optischen Systems
(ρ, ψ)	Ortsfrequenzen in Polarkoordinaten
Ψ_q	Skalierungswert für die Bildqualität
Ψ_s	Skalierungswert für die Bildschärfe
σ_R	Ausdehnungsparameter des visuellen Systems
M_2	zweites Moment der Punktantwort
p'	Position der Kamera
p'_{opt}	beste Fokussierungsposition der Kamera

VAR	Varianz der Grauwerte
m_g	Mittelwert der Grauwerte
M	Anzahl der Bildpunkte vom Bild
Z	Zufallsvariable
p_g	Wahrscheinlichkeit des Grauwerts g
$g(\tilde{Z})$	Wahrscheinlichkeitsfunktion
$p(g)$	Histogramm der Grauwerte
ENT	Entropie der Grauwerte
SMD	"Sum Modulus Difference" der Grauwerte
LS	Leistungsspektrum des Bildes
SL	Signalleistung der Grauwerte
TV	"Thresholded Video Signal Pixel Count" vom Bild
LO	Laplace-Operator der Grauwerte
DVAR	Differential der Varianz
FF	Fokussierungsfunktion
SNR	Signal-zu-Rauschverhältnis
λ	Wellenlänge der Strahlung
$z(x,y)$	Rauschen in der Position (x,y)
$\tilde{\sigma}$	Streuung des Rauschens
(X',Y',Z')	Weltkoordinatensystem
(X,Y,Z)	Kamerakoordinatensystem
I	Translationsvektor
A	gesamte Transformationsmatrix
$\tilde{\alpha}$	Rotationswinkel um die X-Achse
$\tilde{\beta}$	Rotationswinkel um die Y-Achse
$\tilde{\theta}$	Rotationswinkel um die Z-Achse
$R_{\tilde{\alpha}}$	Rotationsmatrix um die X-Achse
$R_{\tilde{\beta}}$	Rotationsmatrix um die Y-Achse
$R_{\tilde{\theta}}$	Rotationsmatrix um die Z-Achse
P	perspektivische Transformationsmatrix
γ	Winkel der Laserstrahlung

l	Abstand zwischen dem Ursprung der Weltkoordinaten und des Lasergeräts
$y_a(t), y'_a(t)$	Istwert des Fokussierungsregelkreises
$y_e(t), y'_e(t)$	Sollwert des Fokussierungsregelkreises
$z(t)$	Störungssignal des Fokussierungsregelkreises
$G_1(s)$	Übertragungsfunktion der Interfacekarte
$G_2(s)$	Übertragungsfunktion des Schrittmotors
$G_3(s)$	Übertragungsfunktion des Spindelschlittensystems
$G_4(s)$	Übertragungsfunktion vom Bildverarbeitungssystem
$G_r(s)$	Übertragungsfunktion vom Regler

10. Literatur

- [1] Hopkins, H.H.:
The frequency response of a defocused optical system.
Proc. Roy. Soc. A, 234, 91-103, 1955.
- [2] Goodman, J.G.:
Effects of Film Nonlinearities on Wavefront Reconstruction
Images of Diffuse Object.
J. Opt. Soc. Am. 55, 1276-1283, 1968.
- [3] Muller, R.A.; Buffington, A.:
Real-time correction of atmospherically degraded telescope images
through image sharpening.
J. Opt. Soc. Am. 64, 1200-1210, 1974.
- [4] Erteza, A.:
Sharpness index and its application to focus control.
Applied Optics, 15, 877-881, 1976.
- [5] Jarvis, R.A.:
Focus Optimisation Criteria for Computer Image Processing.
The microscope 24, 163-180, 1976.
- [6] Boseck, S.:
Ermittlung des optimalen Bildes aus einer photographischen Fokusserie in
der Licht- und Elektronenmikroskopie mit Hilfe der Fraunhoferschen
Beugung (Strukturspektroskopie).
Phot. Korr. 107, 19-31, 1971.
- [7] Linfoot, E.H.:
Information Theory and Optical Images.
J. Opt. Soc. Am. 45, 808-819, 1955.
- [8] Lighthart, G.; Groen, C.A.:
A Comparison of Different Autofocus Algorithms.
IEEE Computer Vision and Pattern Recognition, 597-600, 1982.
- [9] Grossmann, P.:
Depth from focus.
Pattern Recognition Letters 5, 63-69, 1987.
- [10] Darrell, T.; Wohn, K.:
Pyramid Based Depth from Focus.
IEEE Computer Vision and Pattern Recognition, 504-509, 1988.
- [11] Pratt, W.K.:
Digital Image Processing.
John Wiley & Sons, New York, 1978.
- [12] Bibermann, L.M.:
Perception of Displayed Information.
Plenum Press NY/London, 1973.
- [13] Pfenninger, U.:
Image Quality Evaluation by Subjective and Objective Criteria of
Sharpness and Gradation with B/W Transparencies.
Diss. ETH No. 7135, 1982.

- [14] Higgins, G.C.:
Image Quality Criteria.
J. App. Phot. Eng. 3, 53-60, 1977.
- [15] Jaroslavskij, L.P.:
Einführung in die Digitale Bildverarbeitung.
VB Deutscher Verlag der Wissenschaften, Berlin, 1985.
- [16] Jähne, B.:
Digitale Bildverarbeitung.
Springer-Verlag, Berlin, 1991.
- [17] Haferkorn, H.:
Optik: Physikalisch-technische Grundlagen und Anwendungen.
Verlag Harri Deutsch, Frankfurt am Main, 1981.
- [18] Schlag, J.F.; Sanderson, A.C.; Neuman, C.P.; Wimberly, F.C.:
Implementation of automatic focusing algorithms for a computer vision system with camera control.
CMU-TI-TR-83-14, Robotics Institute, Carnegie-Mellon University, 1983.
- [19] Ruocco, S.R.:
Sensoren und Wandler für Roboter.
VCH-Verlag, Weinheim, 1991.
- [20] Kazmierczak, H.:
Erfassung und maschinelle Verarbeitung von Bilddaten.
Springer-Verlag, Wien, 1980.
- [21] Falk, D.S.; Brill, D.R.; Stork, D.G.:
Ein Blick ins Licht.
Springer-Verlag, München, 1986.
- [22] Ballard, D.H., Brown, C.M.:
Computer Vision.
Prentice-Hall, New Jersey, 1982.
- [23] Tenenbaum, J.M.:
A Laboratory for Hand-Eye Research.
In: IFIPS, 206-210, 1971.
- [24] Feldman, J.A.:
The Stanford Hand-Eye Project.
In: IJCAII, 521-526, 1969.
- [25] Jarvis, R.A.:
A Perspective on Range Finding Techniques for Computer Vision.
IEEE Transactions on Pattern Analysis and Machine Intelligence,
Vol. PAMI-5, No. 2, 122-139, 1983.
- [26] Jarvis, R.A.:
A laser time-of-flight range scanner for robotic vision.
Australian Nat. Univ., Comput. Sci.Tech.Rep. TR-CS-81-10, 1978.
- [27] Goldberg, N.:
Inside Autofocus: How the Magic works.
Popular Photography, 77-83, February, 1982.

- [28] Tsai, R.Y.:
An efficient and accurate camera calibration technique for 3D machine vision.
Proc. CVPR, 364-373, 1986.
- [29] Ganapathy, S.:
Decomposition of transformation metrices for robot vision.
Proc. Int. Conf. Robotic, 130-139, 1984.
- [30] Grosky, W.I.; Tamburino, L.A.:
A unified approach to the linear camera calibration problem.
Proc. ICCU, 511-515, 1987.
- [31] Wong, K.W.:
Mathematical formulation and digital analysis in close-range photogrammetry.
Photogrammetric Engineering and Remote Sensing, Vol. 41, 1355-1373, 1975.
- [32] Weng, J.; Cohen, P.; Herniou, M.:
Calibaiton of stereo cameras using a non-linear distortion model.
Proc. ICPR, 246-253, 1990.
- [33] Martins, H.A.; Birk, J.R.; Kelley, R.B.:
Camera models based on data from two calibration planes.
CGIP 17, 173-180, 1981.
- [34] Nobel, B.:
Applied linear algebra.
Prentice-Hall, Englewood Cliffs, N.J., 1968.
- [35] Ishii, M.; Nagata, T.:
Feature Extraction of 3-Dimensional Objects with a Laser Tracker.
Trans. of the Society of Instrument and Control Engineers, 599-605, 10, 5.1974.
- [36] Giloi, W.; Herschel, R.:
Rechenanleitung für Analogrechner.
AEG-Fachbuch, Konstanz, 1975.
- [37] Ammon, W.:
Schaltungen der Analogrechentechnik.
R. Oldenbourg Verlag, München-Wien, 1966.
- [38] Stute, G.:
Regelung an Werkzeugmaschinen.
Carl Hanser Verlag, München-Wien, 1981.
- [39] Snyder, W.E.:
Computer gesteuerte Industrieroboter. Grundlagen und Einsatz.
VCH Verlag, Weinheim, 1990.
- [40] Dörrscheidt, F.; Latzel, W.:
Grundlagen der Regelungstechnik.
B.G. Teubner, Stuttgart, 1989.

- [41] Gassmann, H.:
Einführung in die Regelungstechnik.
Verlag Harri Deutsch, Thun und Frankfurt am Main, 1986.
- [42] Lunderstädt, R.:
Regelungstechnik 2, Vorlesungsmanuskript.
Universität der Bundeswehr Hamburg, Hamburg, 1988.
- [43] Föllinger, O.:
Regelungstechnik.
Elitera-Verlag, Berlin, 1978.
- [44] Schmidt, G.:
Simulationstechnik.
R. Oldenbourg Verlag, München-Wien, 1980.

-171-

Lebenslauf

Name: Bingzi Liao
Geburtstag: 01.06.1962
Geburtsort: Hubei Provinz, V.R. China
Familienstand: verheiratet

Schulbildung:
03.1969 - 07.1974 Grundschule Yongjiou
09.1974 - 07.1978 Mittelschule Heshengqiao
Abschluß: Abitur

Studium:
10.1978 - 07.1982 Wuhan University of Water Transportation Engineering
Department of Automation and Computer Science
Fachrichtung Automatisierungstechnik
Abschluß: Bachelor - Ing.

09.1982 - 03.1985 Aspirantenabteilung of Wuhan University of
Water Transportation Engineering
Fachrichtung Automatisierungstechnik
Abschluß: Magister - Ing.

09.1987 - 07.1988 Besuch des Deutschkurses an der Fremdsprachen-
Universität Shanghai

04.1989 - 09.1989 Besuch des Deutschkurses des Goethe-Instituts
Freiburg

Berufstätigkeit:
05.1985 - 04.1989 Dozent an dem Department of Automation and
Computer Science, Wuhan University of Water
Transportation Engineering

seit 09.1989 Gastwissenschaftler am Institut für
Automatisierungstechnik an der Universität
der Bundeswehr Hamburg

